GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2005, 09:08:01 ; Search time 158 Seconds
(without alignments)
41.613 Million cell updates/sec

Title: US-09-845-736-1
Perfect score: 88
Sequence: 1 SSKITHRIHWESASLLR 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	17	6	ABU08840	Abu08840 Complemen
2	88	100.0	17	6	ABU08842	Abu08842 Complemen
3	88	100.0	17	6	ADA15714	Ada15714 Human bio
4	88	100.0	17	6	ABB82769	Abb82769 Congestiv
5	88	100.0	17	6	ABU08806	Abu08806 C3f compl
6	88	100.0	17	6	ABU08617	Abu08617 Disease s
7	88	100.0	17	8	ADQ96607	Adq96607 Human C3
8	88	100.0	17	8	ADR49407	Adr49407 Autism re
9	88	100.0	18	6	ABU08618	Abu08618 Disease s
10	88	100.0	705	7	ADD93520	Add93520 Novel NOV
11	88	100.0	1255	6	ABR63374	Abr63374 Human Alz
12	88	100.0	1540	4	ABG25976	Abg25976 Novel hum
13	88	100.0	1592	2	AAW34623	Aaw34623 Human C3
14	88	100.0	1635	2	AAW34624	Aaw34624 Human C3
15	88	100.0	1657	2	AAW34629	Aaw34629 Human C3
16	88	100.0	1661	2	AAW34625	Aaw34625 Human C3
17	88	100.0	1663	2	AAR94029	Aar94029 Human mod
18	88	100.0	1663	2	AAR94030	Aar94030 Human mod
19	88	100.0	1663	2	AAR94028	Aar94028 Human C3
20	88	100.0	1663	2	AAW34611	Aaw34611 Human C3
21	88	100.0	1663	2	AAW34621	Aaw34621 Human C3
22	88	100.0	1663	2	AAW40990	Aaw40990 Human C3
23	88	100.0	1663	2	AAW34619	Aaw34619 Human C3
24	88	100.0	1663	2	AAW34617	Aaw34617 Human C3
25	88	100.0	1663	2	AAW34628	Aaw34628 Human C3
26	88	100.0	1663	2	AAW34607	Aaw34607 Human C3
27	88	100.0	1663	2	AAW34606	Aaw34606 Wild type
28	88	100.0	1663	2	AAW34610	Aaw34610 Human C3
29	88	100.0	1663	2	AAW34614	Aaw34614 Human C3
30	88	100.0	1663	2	AAW34616	Aaw34616 Human C3
31	88	100.0	1663	2	AAW34613	Aaw34613 Human C3
32	88	100.0	1663	2	AAW34620	Aaw34620 Human C3
33	88	100.0	1663	2	AAW34627	Aaw34627 Human C3
34	88	100.0	1663	2	AAW34630	Aaw34630 Human C3
35	88	100.0	1663	2	AAW34618	Aaw34618 Human C3
36	88	100.0	1663	2	AAW34612	Aaw34612 Human C3
37	88	100.0	1663	2	AAW34615	Aaw34615 Human C3
38	88	100.0	1663	2	AAW40989	Aaw40989 Human C3
39	88	100.0	1663	7	ADB90023	Adb90023 House com
40	88	100.0	1663	7	ADD93518	Add93518 Novel NOV
41	88	100.0	1663	8	ADK12322	Adk12322 Human com
42	88	100.0	1663	8	ADN04780	Adn04780 Antipsori
43	88	100.0	1663	8	ADP24810	Adp24810 PRO polyp
44	88	100.0	1663	8	ADQ39664	Adq39664 Human myo
45	88	100.0	1667	2	AAW34626	Aaw34626 Human C3

ALIGNMENTS

RESULT 1
ABU08840
ID ABU08840 standard; peptide; 17 AA.
XX
AC ABU08840;
XX
DT 25-AUG-2003 (first entry)
XX
DE Complement C3f peptide, #6, used for physiological condition diagnostics.
XX
KW Proteomic; human; physiological condition; analyte; biopolymer;
KW biomarker; complement C3f; intracerebral haemorrhage; ICH; CHF;
KW congestive heart failure; myocardial infarction; MI; stroke;
KW type II diabetes.
XX
OS Homo sapiens.
XX
PN US2002160420-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00846330.
XX
PR 30-APR-2001; 2001US-00846330.
XX
PA (JACK/) JACKOWSKI G.
PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-491923/46.
XX
PT Determining proteomic basis e.g. basis for diagnosing existence of or
PT predicting development and/or progression of abnormal physiological
PT conditions based upon the presence of proteomic materials.
XX
PS Disclosure; Page 17; 25pp; English.
XX
CC The invention discloses a method for determining a proteomic basis for
CC development and progression of abnormal physiological conditions. The
CC method comprises isolating one or more patient specific proteomic
CC materials from a sample and comparing it against a library of proteomic
CC materials having characteristics identifiable with both normal and
CC abnormal physiological conditions or their predictive hallmarks. The
CC method is useful for determining a proteomic basis for development and

CC progression of abnormal physiological conditions. The method is also
CC useful for evaluating samples containing several analytes/biopolymers for
CC the presence of physiological condition specific sequences. The peptide
CC presented is a biomarker from complement C3f and is associated with
CC intracerebral haemorrhage (ICH), congestive heart failure (CHF),
CC myocardial infarction (MI), type II diabetes and stroke
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 2
ABU08842
ID ABU08842 standard; peptide; 17 AA.
XX
AC ABU08842;
XX
DT 25-AUG-2003 (first entry)
XX
DE Complement C3f peptide, #7, used for physiological condition diagnostics.
XX
KW Proteomic; human; physiological condition; analyte; biopolymer;
KW biomarker; complement C3f; intracerebral haemorrhage; ICH; CHF;
KW congestive heart failure; myocardial infarction; MI; stroke;
KW type II diabetes.
XX
OS Homo sapiens.
XX
PN US2002160420-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00846330.
XX
PR 30-APR-2001; 2001US-00846330.
XX
PA (JACK/) JACKOWSKI G.
PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-491923/46.
XX
PT Determining proteomic basis e.g. basis for diagnosing existence of or
PT predicting development and/or progression of abnormal physiological
PT conditions based upon the presence of proteomic materials.
XX
PS Disclosure; Page 18; 25pp; English.
XX
CC The invention discloses a method for determining a proteomic basis for
CC development and progression of abnormal physiological conditions. The
CC method comprises isolating one or more patient specific proteomic
CC materials from a sample and comparing it against a library of proteomic
CC materials having characteristics identifiable with both normal and
CC abnormal physiological conditions or their predictive hallmarks. The
CC method is useful for determining a proteomic basis for development and
CC progression of abnormal physiological conditions. The method is also
CC useful for evaluating samples containing several analytes/biopolymers for
CC the presence of physiological condition specific sequences. The peptide
CC presented is a biomarker from complement C3f and is associated with
CC intracerebral haemorrhage (ICH), congestive heart failure (CHF),
CC myocardial infarction (MI), type II diabetes and stroke
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 3
ADA15714
ID ADA15714 standard; peptide; 17 AA.
XX
AC ADA15714;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human biopolymer indicative of a disease state.
XX
KW Biopolymer marker; C3f; complement system; myocardial infarction; MI;
KW intracerebral haemorrhage; ICH; congestive heart failure; CHF;
KW type II diabetes; kidney failure; heart failure; Syndrome X; stroke;
KW human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "Optionally absent"
FT Misc-difference 17
FT /note= "Optionally absent"
XX
PN US2002160434-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00845735.
XX
PR 30-APR-2001; 2001US-00845735.
XX
PA (JACK/) JACKOWSKI G.
PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-219988/21.
XX
PT Novel biopolymer marker useful in indicating at least one particular
PT disease state e.g. myocardial infarction, intracerebral hemorrhage,
PT congestive heart failure and type II diabetes.
XX
PS Claim 1; Page 7; 10pp; English.
XX
CC The invention discloses a biopolymer marker useful in indicating at least
CC one particular disease state. The marker is characterised as a C3f
CC fragment from the complement system and is useful for indicating at least
CC one particular disease state e.g. myocardial infarction (MI),
CC intracerebral haemorrhage (ICH), congestive heart failure (CHF) and type
CC II diabetes. Promulgation of various forms of risk assessment tests are
CC contemplated using the biopolymer marker, to identify asymptomatic
CC patients before they suffer an irreversible event such as diabetes,
CC kidney failure and heart failure, and enable effective disease management
CC and preventative medicine. Additionally, the specific diagnostic tests
CC which evolve using the biopolymer marker provide a tool for rapidly and
CC accurately diagnosing acute Syndrome X such as heart attack and stroke,
CC and facilitate treatment. The sequence presented is the biopolymer of the
CC invention.
XX

SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 4
ABB82769
ID ABB82769 standard; peptide; 17 AA.
XX
AC ABB82769;
XX
DT 18-MAR-2003 (first entry)
XX
DE Congestive heart failure indicative biopolymer marker.
XX
KW Biopolymer; marker; C3f; complement system; congestive heart failure;
KW human.
XX
OS Homo sapiens.
XX
PN WO200288717-A2.
XX
PD 07-NOV-2002.
XX
PF 25-APR-2002; 2002WO-CA000578.
XX
PR 30-APR-2001; 2001US-00845736.
XX
PA (SYNX-) SYN.X PHARMA INC.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-120486/11.
XX
PT Use of biopolymer marker for evidencing, categorizing or regulating at
PT least one disease state, e.g. congestive heart failure.
XX
PS Claim 1; Fig 1; 27pp; English.
XX
CC The present sequence represents a biopolymer marker of the invention and
CC is a disease specific marker. The marker is characterised as a C3f
CC fragment from the complement system having a molecular weight of about
CC 2056 daltons. The biopolymer marker identified is useful for evidencing,
CC categorizing or regulating at least one disease state, preferably
CC congestive heart failure
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 5
ABU08806
ID ABU08806 standard; peptide; 17 AA.
XX
AC ABU08806;
XX
DT 20-JUN-2003 (first entry)
XX
DE C3f complement system fragment biopolymer.
XX
KW Biopolymer; complement system; C3f; myocardial infarction;
KW congestive heart failure; Syndrome-X; Type II diabetes.
XX
OS Unidentified.
XX
PN US2002161181-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00846344.
XX
PR 30-APR-2001; 2001US-00846344.
XX
PA (JACK/) JACKOWSKI G.
PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-340873/32.
XX
PT Biopolymer marker, useful in diagnosing disease states including
PT myocardial infarction and Type II diabetes, comprises a compliment C3f
PT fragment with a specified molecular weight.
XX
PS Claim 1; Page 7; 10pp; English.
XX
CC The invention discloses a biopolymer marker which comprises a complement
CC C3f fragment. The marker is used in methods for diagnosing disease states
CC including myocardial infarction, congestive heart failure, Syndrome-X
CC and/or Type II diabetes. The methods used include mass spectroscopy or
CC immunoassays, e.g. radioimmunoassay, enzyme-linked immunosorbent assay
CC (ELISA) or fluorescent immunoassays. The invention enables the
CC characterisation of the presence or absence of the disease state relative
CC to recognition of the presence or absence of the biopolymer,
CC respectively. The sequence presented is the biopolymer of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 6
ABU08617
ID ABU08617 standard; peptide; 17 AA.
XX
AC ABU08617;
XX
DT 23-MAY-2003 (first entry)
XX
DE Disease specific biopolymer marker #1.
XX
KW Biopolymer marker; type II diabetes; immunoassay.
XX
OS Homo sapiens.
XX
PN US2002160532-A1.
XX
PD 31-OCT-2002.
XX
PF 30-APR-2001; 2001US-00846346.
XX
PR 30-APR-2001; 2001US-00846346.
XX
PA (JACK/) JACKOWSKI G.

PA (THAT/) THATCHER B.
PA (MARS/) MARSHALL J.
PA (YANT/) YANTHA J.
PA (VREE/) VREES T.
XX
PI Jackowski G, Thatcher B, Marshall J, Yantha J, Vrees T;
XX
DR WPI; 2003-328370/31.
XX
PT Biopolymer marker useful in indicating disease state, in particular type
PT II diabetes and as antigens in immunoassays for detecting individuals
PT suffering from disease known to be evidenced by marker sequence.
XX
PS Claim 1; Page 7; 10pp; English.
XX
CC The invention describes a biopolymer marker (I) useful in indicating at
CC least one particular disease state. (I) is useful for indicating a
CC disease state, in particular type II diabetes. The marker sequences are
CC useful as antigens in immunoassays for the detection of those individuals
CC suffering from the disease known to be evidenced by the marker sequence.
CC (I) provides an efficient diagnostic tool for rapidly and accurately
CC diagnosing disease states such as type II diabetes. This is the amino
CC acid sequence of a biopolymer marker
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 7
ADQ96607
ID ADQ96607 standard; peptide; 17 AA.
XX
AC ADQ96607;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human C3 peptide seqid 9.
XX
KW myocardial infarction; serum protein profile; signal intensity;
KW reference serum protein profile; mass spectrometry;
KW post-translational modification; major serum protein; C3; human.
XX
OS Homo sapiens.
XX
PN US2004121306-A1.
XX
PD 24-JUN-2004.
XX
PF 20-DEC-2002; 2002US-00325162.
XX
PR 20-DEC-2002; 2002US-00325162.
XX
PA (KUPC/) KUPCHAK P.
PA (JACK/) JACKOWSKI G.
PA (MARS/) MARSHALL J.
XX
PI Kupchak P, Jackowski G, Marshall J;
XX
DR WPI; 2004-532671/51.
XX
PT Diagnosing and distinguishing myocardial infarction in human, involves
PT comparing serum protein profile of human to reference serum protein
PT profiles of at least two subsets of human.
XX
PS Disclosure; SEQ ID NO 9; 33pp; English.
XX
CC The invention describes a method of diagnosing and distinguishing
CC myocardial infarction in a human. The method involves: identifying areas
CC of the serum protein profiles that are different in signal intensity
CC between the human and defined subsets of human; reducing the
CC dimensionality of the areas identified so that the signal intensities
CC associated with protein masses identified are retained for each
CC particular subject; elucidating a metric to identify a human subject;
CC comparing a serum protein profile of the human to the reference serum
CC protein profiles; and analysing the metric elucidated to identify a human
CC subject based upon statistical comparison of characteristics of the
CC reference serum protein profiles of human subjects. The difference in
CC signal intensity represents a difference in protein mass. The serum
CC protein profiles are generated by mass spectrometry. The method is used
CC for diagnosing and distinguishing myocardial infarction in human for use
CC in e.g. post-translational modifications of major serum proteins. The
CC method is simple, economical and does require specialised reagents or
CC optimised assays. This is the amino acid sequence of a C3 peptide
CC identification of which is associated with the diagnosis of myocardial
CC infarction.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 8
ADR49407
ID ADR49407 standard; peptide; 17 AA.
XX
AC ADR49407;
XX
DT 02-DEC-2004 (first entry)
XX
DE Autism related peptide, seq id 3.
XX
KW Autism; marker; peptide; tissue sample analysis.
XX
OS Homo sapiens.
XX
PN WO2004079371-A1.
XX
PD 16-SEP-2004.
XX
PF 02-MAR-2004; 2004WO-SE000193.
XX
PR 04-MAR-2003; 2003SE-00000586.
XX
PA (FORS-) FORSKARPATENT I SYD AB.
XX
PI Grubb A;
XX
DR WPI; 2004-662478/64.
XX
PT Diagnosing autism in subjects suspected of suffering of autism comprises
PT analyzing a tissue, body liquid and/or plasma sample for the presence of
PT high concentrations of certain peptides having specific molecular masses.
XX
PS Claim 2; SEQ ID NO 3; 14pp; English.
XX
CC The invention relates to diagnosing autism in subjects suspected of
CC suffering of autism. The method comprises analysing a tissue sample, a
CC body liquid sample and/or a plasma sample with regard to the presence of
CC high concentrations of certain peptides having the molecular masses 1779
CC +/- 1 Da, 1865 +/- 1 Da and 2022 +/- 1 Da, respectively. Further
CC disclosed is a kit comprising a marker for certain peptides defined
CC above. The amount of the peptide(s) is more than 10 times that present in
CC a sample of a non-autistic subject. Any quantitative immunochemical

```
CC   method is used. The respective peptide is determined using enzyme-linked
CC   immunosorbant assay technology, RIA technology, or SELDI-TOF-MS system.
CC   The method is useful for diagnosing autism in subjects suspected of
CC   suffering of autism, and thus might lead to new possibilities of
CC   treatment of autism, e.g., by suppressing the formation of these
CC   peptides. The current sequence represents a peptide of the invention.
XX
SQ   Sequence 17 AA;

  Query Match             100.0%;  Score 88;  DB 8;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db          1 SSKITHRIHWESASLLR 17


RESULT 9
ABU08618
ID   ABU08618 standard; peptide; 18 AA.
XX
AC   ABU08618;
XX
DT   23-MAY-2003  (first entry)
XX
DE   Disease specific biopolymer marker #2.
XX
KW   Biopolymer marker; type II diabetes; immunoassay.
XX
OS   Homo sapiens.
XX
PN   US2002160532-A1.
XX
PD   31-OCT-2002.
XX
PF   30-APR-2001; 2001US-00846346.
XX
PR   30-APR-2001; 2001US-00846346.
XX
PA   (JACK/) JACKOWSKI G.
PA   (THAT/) THATCHER B.
PA   (MARS/) MARSHALL J.
PA   (YANT/) YANTHA J.
PA   (VREE/) VREES T.
XX
PI   Jackowski G,  Thatcher B,  Marshall J,  Yantha J,  Vrees T;
XX
DR   WPI; 2003-328370/31.
XX
PT   Biopolymer marker useful in indicating disease state, in particular type
PT   II diabetes and as antigens in immunoassays for detecting individuals
PT   suffering from disease known to be evidenced by marker sequence.
XX
PS   Disclosure; Fig 2; 10pp; English.
XX
CC   The invention describes a biopolymer marker (I) useful in indicating at
CC   least one particular disease state. (I) is useful for indicating a
CC   disease state, in particular type II diabetes. The marker sequences are
CC   useful as antigens in immunoassays for the detection of those individuals
CC   suffering from the disease known to be evidenced by the marker sequence.
CC   (I) provides an efficient diagnostic tool for rapidly and accurately
CC   diagnosing disease states such as type II diabetes. This is the amino
CC   acid sequence of a biopolymer marker
XX
SQ   Sequence 18 AA;

  Query Match             100.0%;  Score 88;  DB 6;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-07;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db          2 SSKITHRIHWESASLLR 18


RESULT 10
ADD93520
ID   ADD93520 standard; protein; 705 AA.
XX
AC   ADD93520;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Novel  NOV1b, homologous to human complement C3 precursor.
XX
KW   NOV1a; human; complement C3; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .22
FT                   /note= "Signal peptide"
FT   Protein         23. .705
FT                   /note= "Mature protein"
XX
PN   WO2003078572-A2.
XX
PD   25-SEP-2003.
XX
PF   06-MAR-2003; 2003WO-US006859.
XX
PR   15-MAR-2002; 2002US-0365034P.
PR   19-MAR-2002; 2002US-0365477P.
PR   21-MAR-2002; 2002US-0366420P.
PR   05-MAR-2003; 2003US-00379747.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Burgess CE,  Chant JS,  Chaudhuri A,  Edinger SR,  Gangolli EA;
PI   Malyankar UM,  Miller CE,  Ooi CE,  Ort T,  Patturajan M,  Rastelli L;
PI   Rieger DK,  Shimkets RA,  Zerhusen BD;
XX
DR   WPI; 2003-779122/73.
DR   N-PSDB; ADD93519.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; Page 101; 205pp; English.
XX
CC   The present sequence is the protein sequence of a novel polypeptide,
CC   designated NOV1b, that shows sequence homology to the human complement C3
CC   precursor. The invention is based on the identification of proteins and
CC   polypeptides, and the nucleic acids encoding them, that are
CC   differentially modulated in a pathological state, disease or an abnormal
CC   condition or state. These are targets for therapeutic agents and can be
CC   used in screening methodologies to identify candidate therapeutic agents
CC   which interact with the target and thereby exert a desired or favourable
CC   effect, e.g. in neurogenesis, cell differentiation, cell proliferation,
CC   haematopoiesis, wound healing and angiogenesis. Methods for diagnosis,
CC   treatment and prevention of disorders involving the novel human nucleic
CC   acids and proteins are provided. The polypeptides can also be used to
CC   raise antibodies useful e.g. in diagnosis and therapy.
XX
SQ   Sequence 705 AA;

  Query Match             100.0%;  Score 88;  DB 7;  Length 705;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db        346 SSKITHRIHWESASLLR 362
```

Same as #6 09/846,346

bad date

RESULT 11
ABR63374
ID ABR63374 standard; protein; 1255 AA.
XX
AC ABR63374;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human Alzheimer's disease associated C3 protein precursor.
XX
KW Alzheimer's disease; human; complement C3 protein precursor;
KW C3f fragment; dementia; MAC3; neuroprotective; nootropic.
XX
OS Homo sapiens.
XX
PN WO2003048775-A2.
XX
PD 12-JUN-2003.
XX
PF 27-NOV-2002; 2002WO-DE004360.
XX
PR 28-NOV-2001; 2001DE-01058180.
XX
PA (BIOV-) BIOVISION AG.
XX
PI Lamping N, Zucht H, Selle H, Juergens M, Heine G, Hess R;
XX
DR WPI; 2003-482818/45.
DR N-PSDB; ACC59422.
XX
PT Detecting Alzheimer's disease or predisposition to it, by detecting
PT altered levels of MAC3 or related peptides, also new peptides and related
PT antibodies and nucleic acids for use in the assay.
XX
PS Disclosure; Page 61-69; 89pp; German.
XX
CC The present invention relates to a method of detecting Alzheimer's
CC disease, or a predisposition to it, by determining the presence of at
CC least one MAC3 marker peptide in a patient sample. This is a variant of
CC the complement C3 protein precursor. The method is used to diagnose
CC Alzheimer's disease, or a predisposition to it, including differentiation
CC from other forms of dementia, also, when made quantitative, for assessing
CC the severity of disease. The marker peptides (also their peptidomimetics,
CC derived antibodies (Ab) and nucleic acid encoding them) are useful
CC therapeutically (modulating the concentration of C3 and MAC3 peptides)
CC and for diagnosis; also to develop new therapies and to monitor
CC treatment. The present sequence is the human complement C3 protein
CC precursor
XX
SQ Sequence 1255 AA;

Query Match 100.0%; Score 88; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 896 SSKITHRIHWESASLLR 912

RESULT 12
ABG25976
ID ABG25976 standard; protein; 1540 AA.
XX
AC ABG25976;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #25967.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS90163.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 56335; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1540 AA;

Query Match 100.0%; Score 88; DB 4; Length 1540;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 13
AAW34623
ID AAW34623 standard; protein; 1592 AA.
XX
AC AAW34623;
XX
DT 09-APR-1998 (first entry)
XX
DE Human C3 protein mutant FT-1.
XX
KW Human; C3 protein; convertase; complement pathway protein; infection;
KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

```
KW   complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1591
FT                   /note= "R1591T mutation"
FT   Misc-difference 1592
FT                   /note= "E1592N mutation"
FT   Misc-difference 1593
FT                   /note= "A1593Stop mutation"
XX
PN   WO9732981-A1.
XX
PD   12-SEP-1997.
XX
PF   04-MAR-1997;   97WO-GB000603.
XX
PR   07-MAR-1996;   96GB-00004865.
PR   07-JUN-1996;   96GB-00011896.
PR   08-JUL-1996;   96GB-00014293.
PR   19-NOV-1996;   96GB-00024028.
XX
PA   (IMUT-) IMUTRAN LTD.
XX
PI   Farries TC,  Harrison RA;
XX
DR   WPI; 1997-457534/42.
XX
PT   Modified complement pathway protein that forms C3 convertase resistant to
PT   down-regulation - used to exhaust the complement pathway by super-
PT   activation, especially for preventing graft rejection, etc.
XX
PS   Example 17; Page; 123pp; English.
XX
CC   This sequence represents a mutated human C3 protein of the invention (see
CC   AAW34606 for wild type protein). This protein is a protein of the
CC   invention, and is a modified native complement pathway protein (A) that
CC   forms a down-regulation resistant C3 convertase. (A), their variants,
CC   fragments and conjugates are used to deplete levels of complement pathway
CC   proteins (by superactivation until one or more components are exhausted),
CC   specifically to prevent rejection of foreign material (particularly a
CC   xenograft) but also to prevent complement-mediated diseases resulting
CC   from (surgical) injury or antibody-antigen interaction in autoimmune
CC   disease, also to localise and/or amplify endogenous complement protein
CC   conversion and deposition at a specific site (e.g. a virus, infected cell
CC   or tumour, to increase sensitivity to complement-mediated responses; a
CC   particular application is eliminating any cancer cells left after
CC   surgical removal of a tumour). Also contemplated is ex vivo treatment,
CC   especially by passing blood through a matrix containing (A) (this may
CC   remove additional anaphylactic peptides and other inflammatory mediators)
CC   or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
CC   bone marrow. Since (A) is not inhibited by factor I, it can bind
CC   repeatedly to factor B (which is then inactivated), causing inactivation
CC   of the alternative pathway by consumption of factor B
XX
SQ   Sequence 1592 AA;

  Query Match           100.0%;  Score 88;  DB 2;  Length 1592;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 14
AAW34624
ID   AAW34624 standard; protein; 1635 AA.
XX
AC   AAW34624;
XX
DT   09-APR-1998  (first entry)
XX
DE   Human C3 protein mutant FT-2.
XX
KW   Human; C3 protein; convertase; complement pathway protein; infection;
KW   down-regulation resistant C3 convertase; xenograft rejection; therapy;
KW   complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
KW   complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1636
FT                   /note= "wild type E mutated to stop codon"
XX
PN   WO9732981-A1.
XX
PD   12-SEP-1997.
XX
PF   04-MAR-1997;   97WO-GB000603.
XX
PR   07-MAR-1996;   96GB-00004865.
PR   07-JUN-1996;   96GB-00011896.
PR   08-JUL-1996;   96GB-00014293.
PR   19-NOV-1996;   96GB-00024028.
XX
PA   (IMUT-) IMUTRAN LTD.
XX
PI   Farries TC,  Harrison RA;
XX
DR   WPI; 1997-457534/42.
XX
PT   Modified complement pathway protein that forms C3 convertase resistant to
PT   down-regulation - used to exhaust the complement pathway by super-
PT   activation, especially for preventing graft rejection, etc.
XX
PS   Example 17; Page; 123pp; English.
XX
CC   This sequence represents a mutated human C3 protein of the invention (see
CC   AAW34606 for wild type protein). This protein is a protein of the
CC   invention, and is a modified native complement pathway protein (A) that
CC   forms a down-regulation resistant C3 convertase. (A), their variants,
CC   fragments and conjugates are used to deplete levels of complement pathway
CC   proteins (by superactivation until one or more components are exhausted),
CC   specifically to prevent rejection of foreign material (particularly a
CC   xenograft) but also to prevent complement-mediated diseases resulting
CC   from (surgical) injury or antibody-antigen interaction in autoimmune
CC   disease, also to localise and/or amplify endogenous complement protein
CC   conversion and deposition at a specific site (e.g. a virus, infected cell
CC   or tumour, to increase sensitivity to complement-mediated responses; a
CC   particular application is eliminating any cancer cells left after
CC   surgical removal of a tumour). Also contemplated is ex vivo treatment,
CC   especially by passing blood through a matrix containing (A) (this may
CC   remove additional anaphylactic peptides and other inflammatory mediators)
CC   or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
CC   bone marrow. Since (A) is not inhibited by factor I, it can bind
CC   repeatedly to factor B (which is then inactivated), causing inactivation
CC   of the alternative pathway by consumption of factor B
XX
SQ   Sequence 1635 AA;

  Query Match           100.0%;  Score 88;  DB 2;  Length 1635;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 15
AAW34629
ID   AAW34629 standard; protein; 1657 AA.
```

```
XX
AC   AAW34629;
XX
DT   09-APR-1998  (first entry)
XX
DE   Human C3 protein mutant FR-2.
XX
KW   Human; C3 protein; convertase; complement pathway protein; infection;
KW   down-regulation resistant C3 convertase; xenograft rejection; therapy;
KW   complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
KW   complement-mediated response; MHC-mismatched lymphocyte; mutein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1638. .1645
FT                   /note= "wild type residues QDEENQKQ mutated to SS"
XX
PN   WO9732981-A1.
XX
PD   12-SEP-1997.
XX
PF   04-MAR-1997;   97WO-GB000603.
XX
PR   07-MAR-1996;   96GB-00004865.
PR   07-JUN-1996;   96GB-00011896.
PR   08-JUL-1996;   96GB-00014293.
PR   19-NOV-1996;   96GB-00024028.
XX
PA   (IMUT-) IMUTRAN LTD.
XX
PI   Farries TC,  Harrison RA;
XX
DR   WPI; 1997-457534/42.
XX
PT   Modified complement pathway protein that forms C3 convertase resistant to
PT   down-regulation - used to exhaust the complement pathway by super-
PT   activation, especially for preventing graft rejection, etc.
XX
PS   Example 17; Page; 123pp; English.
XX
CC   This sequence represents a mutated human C3 protein of the invention (see
CC   AAW34606 for wild type protein). This protein is a protein of the
CC   invention, and is a modified native complement pathway protein (A) that
CC   forms a down-regulation resistant C3 convertase. (A), their variants,
CC   fragments and conjugates are used to deplete levels of complement pathway
CC   proteins (by superactivation until one or more components are exhausted),
CC   specifically to prevent rejection of foreign material (particularly a
CC   xenograft) but also to prevent complement-mediated diseases resulting
CC   from (surgical) injury or antibody-antigen interaction in autoimmune
CC   disease, also to localise and/or amplify endogenous complement protein
CC   conversion and deposition at a specific site (e.g. a virus, infected cell
CC   or tumour, to increase sensitivity to complement-mediated responses; a
CC   particular application is eliminating any cancer cells left after
CC   surgical removal of a tumour). Also contemplated is ex vivo treatment,
CC   especially by passing blood through a matrix containing (A) (this may
CC   remove additional anaphylactic peptides and other inflammatory mediators)
CC   or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
CC   bone marrow. Since (A) is not inhibited by factor I, it can bind
CC   repeatedly to factor B (which is then inactivated), causing inactivation
CC   of the alternative pathway by consumption of factor B
XX
SQ   Sequence 1657 AA;

  Query Match             100.0%;  Score 88;  DB 2;  Length 1657;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320
```

WO

```
Search completed: June  1, 2005, 09:31:13
Job time : 160 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2005, 09:14:16 ; Search time 40 Seconds
(without alignments)
31.726 Million cell updates/sec

Title: US-09-845-736-1
Perfect score: 88
Sequence: 1 SSKITHRIHWESASLLR 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	17	4	US-09-846-345A-1	Sequence 1, Appli
2	88	100.0	17	4	US-09-846-344-1	Sequence 1, Appli
3	88	100.0	1663	2	US-08-793-126-1	Sequence 1, Appli
4	88	100.0	1663	3	US-09-132-271-1	Sequence 1, Appli
5	88	100.0	1663	3	US-09-142-334-22	Sequence 22, Appl
6	84	95.5	16	4	US-09-845-730A-1	Sequence 1, Appli
7	75	85.2	14	4	US-09-846-349A-1	Sequence 1, Appli
8	52	59.1	281	4	US-09-252-991A-22644	Sequence 22644, A
9	45	51.1	280	4	US-09-252-991A-21635	Sequence 21635, A
10	44	50.0	146	4	US-09-640-211A-2215	Sequence 2215, Ap
11	41	46.6	221	4	US-09-252-991A-24616	Sequence 24616, A
12	39.5	44.9	433	4	US-09-328-352-7646	Sequence 7646, Ap
13	39	44.3	264	4	US-09-107-532A-5290	Sequence 5290, Ap
14	39	44.3	329	4	US-09-107-532A-7038	Sequence 7038, Ap
15	39	44.3	438	4	US-09-134-000C-4100	Sequence 4100, Ap
16	39	44.3	817	4	US-09-543-681A-4637	Sequence 4637, Ap
17	39	44.3	1060	4	US-09-489-039A-11403	Sequence 11403, A
18	39	44.3	1525	4	US-09-418-710-69	Sequence 69, Appl
19	39	44.3	1525	4	US-09-839-479-68	Sequence 68, Appl
20	39	44.3	1527	4	US-09-418-710-27	Sequence 27, Appl
21	39	44.3	1527	4	US-09-839-479-27	Sequence 27, Appl
22	39	44.3	1531	4	US-09-418-710-29	Sequence 29, Appl
23	39	44.3	1531	4	US-09-839-479-29	Sequence 29, Appl
24	39	44.3	1540	4	US-09-949-016-7037	Sequence 7037, Ap
25	38	43.2	133	4	US-09-673-395A-177	Sequence 177, App
26	38	43.2	151	4	US-09-270-767-45947	Sequence 45947, A
27	38	43.2	239	4	US-09-270-767-36399	Sequence 36399, A
28	38	43.2	239	4	US-09-270-767-51616	Sequence 51616, A
29	38	43.2	277	4	US-09-489-039A-7268	Sequence 7268, Ap
30	38	43.2	280	4	US-09-399-081A-2	Sequence 2, Appli
31	38	43.2	283	4	US-09-949-016-7398	Sequence 7398, Ap
32	38	43.2	299	4	US-09-640-211A-636	Sequence 636, App
33	38	43.2	313	3	US-09-124-758-4	Sequence 4, Appli
34	38	43.2	313	4	US-09-709-677-4	Sequence 4, Appli
35	38	43.2	313	4	US-09-646-693-2	Sequence 2, Appli
36	38	43.2	313	4	US-09-270-767-43932	Sequence 43932, A
37	38	43.2	316	4	US-09-489-039A-11945	Sequence 11945, A
38	38	43.2	412	4	US-09-252-991A-18569	Sequence 18569, A
39	38	43.2	528	4	US-09-270-767-42895	Sequence 42895, A
40	37	42.0	130	4	US-09-270-767-38113	Sequence 38113, A
41	37	42.0	130	4	US-09-270-767-53330	Sequence 53330, A
42	37	42.0	150	4	US-09-636-215-707	Sequence 707, App
43	37	42.0	150	4	US-09-685-166A-707	Sequence 707, App
44	37	42.0	150	4	US-09-679-426-707	Sequence 707, App
45	37	42.0	150	4	US-09-759-143-707	Sequence 707, App

ALIGNMENTS

RESULT 1
US-09-846-345A-1
; Sequence 1, Application US/09846345A
; Patent No. 6617308
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 1865 DALTONS
; FILE REFERENCE: 2132.045
; CURRENT APPLICATION NUMBER: US/09/846,345A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-345A-1

Query Match 100.0%; Score 88; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 2
US-09-846-344-1
; Sequence 1, Application US/09846344
; Patent No. 6756476
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 2021 DALTONS
; FILE REFERENCE: 2132.048
; CURRENT APPLICATION NUMBER: US/09/846,344
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-344-1

Query Match 100.0%; Score 88; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db          1 SSKITHRIHWESASLLR 17


RESULT 3
US-08-793-126-1
; Sequence 1, Application US/08793126
; Patent No. 5849297
;  GENERAL INFORMATION:
;    APPLICANT:  Harrison, Richard Alexander
;    APPLICANT:  Farries, Charles Timothy
;    TITLE OF INVENTION:  MODIFIED HUMAN C3 PROTEINS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HALE AND DORR LLP
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  United States of America
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,126
;      FILING DATE:  07-FEB-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baker, Hollie L.
;      REGISTRATION NUMBER:  31,321
;      REFERENCE/DOCKET NUMBER:  102286.377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 526-6000
;      TELEFAX:  (617) 526-5000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1663 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-793-126-1

  Query Match           100.0%;  Score 88;  DB 2;  Length 1663;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 4
US-09-132-271-1
; Sequence 1, Application US/09132271
; Patent No. 6221657
;  GENERAL INFORMATION:
;    APPLICANT:  Harrison, Richard Alexander
;    APPLICANT:  Farries, Charles Timothy
;    TITLE OF INVENTION:  MODIFIED HUMAN C3 PROTEINS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HALE AND DORR LLP
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  United States of America
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/132,271
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,126
;      FILING DATE:  07-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baker, Hollie L.
;      REGISTRATION NUMBER:  31,321
;      REFERENCE/DOCKET NUMBER:  102286.377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 526-6000
;      TELEFAX:  (617) 526-5000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1663 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-132-271-1

  Query Match           100.0%;  Score 88;  DB 3;  Length 1663;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 5
US-09-142-334-22
; Sequence 22, Application US/09142334
; Patent No. 6268485
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT: Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/142,334
;  CURRENT FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: PCT/GB97/00603
;  EARLIER FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;    LENGTH: 1663
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-142-334-22

  Query Match           100.0%;  Score 88;  DB 3;  Length 1663;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-05;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 6
US-09-845-730A-1
; Sequence 1, Application US/09845730A
; Patent No. 6593298
; GENERAL INFORMATION:
;  APPLICANT: Jackowski, George
;  TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
```

```
;  TITLE OF INVENTION:  OF 1690 DALTONS
;  FILE REFERENCE: 2132.042
;  CURRENT APPLICATION NUMBER: US/09/845,730A
;  CURRENT FILING DATE:  2001-04-30
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-845-730A-1

  Query Match             95.5%;  Score 84;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLLR 17
              ||||||||||||||||
Db          1 SKITHRIHWESASLLR 16


RESULT 7
US-09-846-349A-1
; Sequence 1, Application US/09846349A
; Patent No. 6602855
; GENERAL INFORMATION:
;  APPLICANT: Jackowski, George
;  APPLICANT:  Vrees, Tammy
;  APPLICANT:  Yantha, Jason
;  APPLICANT:  Marshall, John
;  APPLICANT:  Thatcher, Brad
;  TITLE OF INVENTION: Biopolymer Marker Indicative Of Disease State Having A Molecular
;  TITLE OF INVENTION:  of 1449 Daltons
;  FILE REFERENCE: 2132.034
;  CURRENT APPLICATION NUMBER: US/09/846,349A
;  CURRENT FILING DATE:  2001-04-30
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-846-349A-1

  Query Match             85.2%;  Score 75;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 ITHRIHWESASLLR 17
              ||||||||||||||
Db          1 ITHRIHWESASLLR 14


RESULT 8
US-09-252-991A-22644
; Sequence 22644, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 22644
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22644

  Query Match             59.1%;  Score 52;  DB 4;  Length 281;
  Best Local Similarity   57.1%;  Pred. No. 1.2;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASL 15
              :|| | :||: |||
Db        132 AKIAHHLHWQHASL 145


RESULT 9
US-09-252-991A-21635
; Sequence 21635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 21635
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21635

  Query Match             51.1%;  Score 45;  DB 4;  Length 280;
  Best Local Similarity   40.0%;  Pred. No. 15;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLL 16
              ::| ||: ||:: |:
Db        119 ARIVHRLDWETSGLM 133


RESULT 10
US-09-640-211A-2215
; Sequence 2215, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2215
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Pinus radiata
US-09-640-211A-2215

  Query Match             50.0%;  Score 44;  DB 4;  Length 146;
  Best Local Similarity   53.3%;  Pred. No. 11;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASL 15
              || ::|   |||| |
Db        126 SSNLSHMAQWESARL 140
```

```
RESULT 11
US-09-252-991A-24616
; Sequence 24616, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 24616
;   LENGTH: 221
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24616

  Query Match             46.6%;  Score 41;  DB 4;  Length 221;
  Best Local Similarity   60.0%;  Pred. No. 52;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLL 16
              |  | ||  | |||:|
Db        139 SSSTPRIAWSSASML 153


RESULT 12
US-09-328-352-7646
; Sequence 7646, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7646
;   LENGTH: 433
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-7646

  Query Match             44.9%;  Score 39.5;  DB 4;  Length 433;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 SKITH-RIHWESASLL 16
              ||: |  |:|: ||:|
Db        373 SKLCHEHIYWDQASVL 388


RESULT 13
US-09-107-532A-5290
; Sequence 5290, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 5290:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 264 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...264
;        SEQUENCE DESCRIPTION: SEQ ID NO: 5290:
US-09-107-532A-5290

  Query Match             44.3%;  Score 39;  DB 4;  Length 264;
  Best Local Similarity   53.3%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          4 ITHRI--HWESASLL 16
              :||||  :| :|:||
Db         86 VTHRIIPNWAAAALL 100


RESULT 14
US-09-107-532A-7038
; Sequence 7038, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
```

```
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 7038:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 329 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...329
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7038:
US-09-107-532A-7038

  Query Match             44.3%;  Score 39;  DB 4;  Length 329;
  Best Local Similarity   58.3%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 THRIHWESASLL 16
              |||::::|| ||
Db        105 THRVNFDSALLL 116


RESULT 15
US-09-134-000C-4100
; Sequence 4100, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4100
;   LENGTH: 438
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-4100

  Query Match             44.3%;  Score 39;  DB 4;  Length 438;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 THRIHW 10
              |||:||
Db        390 THRLHW 395


Search completed: June  1, 2005, 09:35:37
Job time : 41 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2005, 09:28:38 ; Search time 139 Seconds
(without alignments)
42.277 Million cell updates/sec

Title: US-09-845-736-1
Perfect score: 88
Sequence: 1 SSKITHRIHWESASLLR 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	17	9	US-09-846-346-1	Sequence 1, Appli
2	88	100.0	17	9	US-09-846-344-1	Sequence 1, Appli
3	88	100.0	17	11	US-09-845-736-1	Sequence 1, Appli
4	88	100.0	17	16	US-10-325-162-9	Sequence 9, Appli
5	88	100.0	17	17	US-10-497-073-1	Sequence 1, Appli
6	88	100.0	705	15	US-10-379-747-4	Sequence 4, Appli
7	88	100.0	1255	17	US-10-497-073-17	Sequence 17, Appl
8	88	100.0	1638	17	US-10-884-813-8	Sequence 8, Appli
9	88	100.0	1638	17	US-10-884-813-12	Sequence 12, Appl
10	88	100.0	1663	9	US-09-875-519A-22	Sequence 22, Appl
11	88	100.0	1663	10	US-09-842-758-41	Sequence 41, Appl
12	88	100.0	1663	15	US-10-379-747-2	Sequence 2, Appli
13	88	100.0	1663	15	US-10-174-333-41	Sequence 41, Appl
14	88	100.0	1663	17	US-10-741-600-1327	Sequence 1327, Ap
15	88	100.0	1663	17	US-10-928-312-2	Sequence 2, Appli
16	88	100.0	1663	17	US-10-884-813-2	Sequence 2, Appli
17	88	100.0	1663	17	US-10-884-813-6	Sequence 6, Appli
18	88	100.0	1663	17	US-10-884-813-10	Sequence 10, Appl
19	84	95.5	16	16	US-10-325-162-10	Sequence 10, Appl
20	84	95.5	16	17	US-10-497-073-7	Sequence 7, Appli
21	83	94.3	16	9	US-09-846-345-1	Sequence 1, Appli
22	83	94.3	16	16	US-10-325-162-8	Sequence 8, Appli
23	83	94.3	16	17	US-10-497-073-2	Sequence 2, Appli
24	80	90.9	15	17	US-10-497-073-9	Sequence 9, Appli
25	79	89.8	15	9	US-09-845-739-1	Sequence 1, Appli
26	79	89.8	15	9	US-09-845-735-1	Sequence 1, Appli
27	79	89.8	15	16	US-10-325-162-7	Sequence 7, Appli
28	79	89.8	15	17	US-10-497-073-8	Sequence 8, Appli
29	75	85.2	14	9	US-09-845-730-1	Sequence 1, Appli
30	75	85.2	14	16	US-10-325-162-6	Sequence 6, Appli
31	75	85.2	14	17	US-10-497-073-3	Sequence 3, Appli
32	75	85.2	14	17	US-10-497-073-10	Sequence 10, Appl
33	75	85.2	14	17	US-10-497-073-11	Sequence 11, Appl
34	70	79.5	13	9	US-09-845-738A-1	Sequence 1, Appli
35	70	79.5	13	16	US-10-325-162-5	Sequence 5, Appli
36	70	79.5	13	17	US-10-497-073-12	Sequence 12, Appl
37	67	76.1	12	17	US-10-497-073-4	Sequence 4, Appli
38	66	75.0	12	9	US-09-846-349-1	Sequence 1, Appli
39	66	75.0	12	16	US-10-325-162-4	Sequence 4, Appli
40	66	75.0	12	17	US-10-497-073-13	Sequence 13, Appl
41	61	69.3	11	9	US-09-845-715-1	Sequence 1, Appli
42	61	69.3	11	16	US-10-325-162-3	Sequence 3, Appli
43	61	69.3	1661	10	US-09-842-758-42	Sequence 42, Appl
44	61	69.3	1661	15	US-10-174-333-42	Sequence 42, Appl
45	58	65.9	10	17	US-10-497-073-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-846-346-1
; Sequence 1, Application US/09846346
; Patent No. US20020160532A1
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 1998 DALTONS
; FILE REFERENCE: 2132.013
; CURRENT APPLICATION NUMBER: US/09/846,346
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-346-1

Query Match 100.0%; Score 88; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 2
US-09-846-344-1
; Sequence 1, Application US/09846344
; Publication No. US20020161181A1
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: BIOPOLYMER MARKER INDICATIVE OF DISEASE STATE HAVING A MOLECULAR
; TITLE OF INVENTION: OF 2021 DALTONS

; FILE REFERENCE: 2132.048
; CURRENT APPLICATION NUMBER: US/09/846,344
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-846-344-1

Query Match 100.0%; Score 88; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 3
US-09-845-736-1
; Sequence 1, Application US/09845736
; Publication No. US20040224423A1
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; APPLICANT: Marshall, John
; APPLICANT: Yantha, Jason
; APPLICANT: Vrees, Tammy
; APPLICANT: Thatcher, Brad
; TITLE OF INVENTION: Biopolymer Marker Indicative of Disease State Having a Molecular
; TITLE OF INVENTION: of 2056 Daltons
; FILE REFERENCE: 2132.049
; CURRENT APPLICATION NUMBER: US/09/845,736
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-845-736-1

Query Match 100.0%; Score 88; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 4
US-10-325-162-9
; Sequence 9, Application US/10325162
; Publication No. US20040121306A1
; GENERAL INFORMATION:
; APPLICANT: Kupchak, Peter
; APPLICANT: Marshall, John
; APPLICANT: Jackowski, George
; TITLE OF INVENTION: Method of Confirming the Presence of Myocardial Infarction
; FILE REFERENCE: 2132.132
; CURRENT APPLICATION NUMBER: US/10/325,162
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-325-162-9

Query Match 100.0%; Score 88; DB 16; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 5
US-10-497-073-1
; Sequence 1, Application US/10497073
; Publication No. US20050048584A1
; GENERAL INFORMATION:
; APPLICANT: BioVisioN AG
; TITLE OF INVENTION: Method for detecting Alzheimer's disease and differentiating
; TITLE OF INVENTION: Alzheimer's disease from other demential diseases, associated
; TITLE OF INVENTION: peptides and the use thereof
; FILE REFERENCE: C3f-PCT
; CURRENT APPLICATION NUMBER: US/10/497,073
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: DE10158180
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: PCT/DE02/04360
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-497-073-1

bad date?

Query Match 100.0%; Score 88; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1 SSKITHRIHWESASLLR 17

RESULT 6
US-10-379-747-4
; Sequence 4, Application US/10379747
; Publication No. US20040023874A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Chant, John S.;
; APPLICANT: Chaudhuri, Amitabha ;
; APPLICANT: Edinger, Shlomit R.;
; APPLICANT: Gangolli, Esha A.;
; APPLICANT: Malyankar, Uriel M.;
; APPLICANT: Miller, Charles E.;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Ort, Tatiana A.;
; APPLICANT: Patturajan, Meera ;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Shimkets, Richard A.;
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568B
; CURRENT APPLICATION NUMBER: US/10/379,747
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/365,034
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/366,420
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/365,477
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 4

; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-747-4

Query Match 100.0%; Score 88; DB 15; Length 705;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 346 SSKITHRIHWESASLLR 362

RESULT 7
US-10-497-073-17
; Sequence 17, Application US/10497073
; Publication No. US20050048584A1
; GENERAL INFORMATION:
; APPLICANT: BioVisioN AG
; TITLE OF INVENTION: Method for detecting Alzheimer's disease and differentiating
; TITLE OF INVENTION: Alzheimer's disease from other demential diseases, associated
; TITLE OF INVENTION: peptides and the use thereof
; FILE REFERENCE: C3f-PCT
; CURRENT APPLICATION NUMBER: US/10/497,073
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: DE10158180
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: PCT/DE02/04360
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-497-073-17

Query Match 100.0%; Score 88; DB 17; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 896 SSKITHRIHWESASLLR 912

RESULT 8
US-10-884-813-8
; Sequence 8, Application US/10884813
; Publication No. US20050079585A1
; GENERAL INFORMATION:
; APPLICANT: Kolln, Johanna
; APPLICANT: Bredehorst, Reinhard
; APPLICANT: Spillner, Edzard
; TITLE OF INVENTION: Complement Depletion with Recombinant Human C3 Derivatives
; FILE REFERENCE: P 63782
; CURRENT APPLICATION NUMBER: US/10/884,813
; CURRENT FILING DATE: 2004-07-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid protein
US-10-884-813-8

Query Match 100.0%; Score 88; DB 17; Length 1638;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 9
US-10-884-813-12
; Sequence 12, Application US/10884813
; Publication No. US20050079585A1
; GENERAL INFORMATION:
; APPLICANT: Kolln, Johanna
; APPLICANT: Bredehorst, Reinhard
; APPLICANT: Spillner, Edzard
; TITLE OF INVENTION: Complement Depletion with Recombinant Human C3 Derivatives
; FILE REFERENCE: P 63782
; CURRENT APPLICATION NUMBER: US/10/884,813
; CURRENT FILING DATE: 2004-07-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid protein
US-10-884-813-12

Query Match 100.0%; Score 88; DB 17; Length 1638;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 10
US-09-875-519A-22
; Sequence 22, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
; APPLICANT: Farries, Timothy C.
; APPLICANT: Harrison, Richard A.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/875,519A
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/GB97/00603
; PRIOR FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-875-519A-22

Query Match 100.0%; Score 88; DB 9; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 11
US-09-842-758-41
; Sequence 41, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Fernandes, Elma R

; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E
; APPLICANT: Gangolli, Esha A
; APPLICANT: Smithson, Glennda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Grosse, William M
; APPLICANT: Edward, Szekeres S
; APPLICANT: Alsobrook II, John P
; TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-758-41

Query Match 100.0%; Score 88; DB 10; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 12
US-10-379-747-2
; Sequence 2, Application US/10379747
; Publication No. US20040023874A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Chant, John S.;
; APPLICANT: Chaudhuri, Amitabha ;
; APPLICANT: Edinger, Shlomit R.;
; APPLICANT: Gangolli, Esha A.;
; APPLICANT: Malyankar, Uriel M.;
; APPLICANT: Miller, Charles E.;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Ort, Tatiana A.;
; APPLICANT: Patturajan, Meera ;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Shimkets, Richard A.;
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568B
; CURRENT APPLICATION NUMBER: US/10/379,747
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: 60/365,034
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/366,420
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/365,477
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-747-2

Query Match 100.0%; Score 88; DB 15; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 13
US-10-174-333-41
; Sequence 41, Application US/10174333
; Publication No. US20040029220A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grosse, William M.
; APPLICANT: Szekeres, Edward S.
; APPLICANT: Alsobrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Li, Li
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-783 CIP1
; CURRENT APPLICATION NUMBER: US/10/174,333

; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/193,664
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/194,614
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,063
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,066
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,067
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,068
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,070
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,510
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/219,855
; PRIOR FILING DATE: 2000-07-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 41
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-333-41

Query Match 100.0%; Score 88; DB 15; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 14
US-10-741-600-1327
; Sequence 1327, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1327
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1327

Query Match 100.0%; Score 88; DB 17; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 15
US-10-928-312-2
; Sequence 2, Application US/10928312
; Publication No. US20050055735A1
; GENERAL INFORMATION:
; APPLICANT: YEUNG SHU-BIU, WILLIAM
; APPLICANT: LEE KAI-FAI, CALVIN
; APPLICANT: LUK MOON-CHING, JOHN
; APPLICANT: LEE YIN LAU, CHERIE
; TITLE OF INVENTION: USE OF COMPLEMENT PROTEIN C3 AND ITS DERIVATIVES IN
; TITLE OF INVENTION: ENHANCING MAMMALIAN EMBRYO DEVELOPMENT
; FILE REFERENCE: V9661.0082
; CURRENT APPLICATION NUMBER: US/10/928,312
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/501,127
; PRIOR FILING DATE: 2003-09-08
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-928-312-2

Query Match 100.0%; Score 88; DB 17; Length 1663;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

Search completed: June 1, 2005, 09:38:01
Job time : 139 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2005, 09:09:26 ; Search time 38 Seconds
(without alignments)
43.044 Million cell updates/sec

Title: US-09-845-736-1
Perfect score: 88
Sequence: 1 SSKITHRIHWESASLLR 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	1663	1	C3HU	complement C3 prec
2	61	69.3	726	2	A27602	complement C3 - ra
3	52	59.1	267	2	A82997	hypothetical prote
4	46	52.3	1663	1	C3RT	complement C3 prec
5	45	51.1	211	2	H83239	pseudouridine synt
6	45	51.1	336	2	F75508	mrr restriction sy
7	45	51.1	1663	1	C3MS	complement C3 prec
8	44	50.0	516	2	S67037	SMP3 protein - yea
9	42	47.7	248	2	AH0011	ferredoxin-NADP re
10	42	47.7	280	2	C86317	protein T10O22.23
11	42	47.7	401	2	E82521	hypothetical prote
12	42	47.7	474	2	G75580	conserved hypothet
13	42	47.7	858	2	T18946	probable phospholi
14	41	46.6	226	1	JQ0393	nodulation protein
15	41	46.6	229	2	AI3289	hypothetical cytos
16	41	46.6	615	2	B86713	hypothetical prote
17	41	46.6	1585	2	AE2916	NAD-glutamate dehy
18	41	46.6	1585	2	H97690	NAD-glutamate dehy
19	41	46.6	1666	1	C3GP	complement C3 prec
20	40.5	46.0	1417	2	H90670	probable invasin [
21	40.5	46.0	1417	2	D85521	probable adhesin e
22	40	45.5	259	2	T29569	hypothetical prote
23	40	45.5	343	2	T42129	probable acyltrans
24	40	45.5	354	2	D41080	probable aldolase
25	40	45.5	593	2	C97848	ABC transporter AT
26	40	45.5	1123	2	T22608	hypothetical prote
27	40	45.5	1456	2	G86466	hypothetical prote
28	40	45.5	2514	2	T37320	ataxia telangiecta
29	40	45.5	2619	2	T24588	hypothetical prote
30	39	44.3	228	2	AI2913	conserved hypothet
31	39	44.3	249	2	T16228	hypothetical prote
32	39	44.3	263	2	T48742	hypothetical prote
33	39	44.3	266	2	D97688	hypothetical prote
34	39	44.3	406	2	T50894	hydroxyneurosporen
35	39	44.3	459	2	B82416	hypothetical prote
36	39	44.3	493	2	G88979	protein F37B4.5 [i
37	39	44.3	567	2	C69611	ABC transporter re
38	39	44.3	574	2	AC1414	ABC transporter re
39	39	44.3	574	2	AB1790	ABC transporter re
40	39	44.3	790	2	S18206	recombination prot
41	39	44.3	851	1	WMBEU9	gene UL9 protein -
42	39	44.3	1015	1	TQECT	transposase - Esch
43	39	44.3	1479	2	T17401	transcription regu
44	39	44.3	1896	2	T08851	Down syndrome cell
45	38	43.2	148	2	A86878	non-heme iron-bind

ALIGNMENTS

RESULT 1
C3HU
complement C3 precursor [validated] - human
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: A94065; A37999; A92187; A27603; A23435; A45830; B45830; A01257; A01258
R;de Bruijn, M.H.L.; Fey, G.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 708-712, 1985
A;Title: Human complement component C3: cDNA coding sequence and derived primary structu
A;Reference number: A94065; MUID:85140166; PMID:2579379
A;Accession: A94065
A;Molecule type: mRNA
A;Residues: 1-1663 <DEB>
A;Cross-references: UNIPROT:P01024; GB:K02765; NID:g179664; PIDN:AAA85332.1; PID:g179665
R;Vik, D.P.; Amiguet, P.; Moffat, G.J.; Fey, M.; Amiguet-Barras, F.; Wetsel, R.A.; Tack,
Biochemistry 30, 1080-1085, 1991
A;Title: Structural features of the human C3 gene: intron/exon organization, transcripti
A;Reference number: A37999; MUID:91113687; PMID:1703437
A;Contents: intron/exon structure of gene
A;Accession: A37999
A;Molecule type: DNA
A;Residues: 1-25 <VIK>
A;Cross-references: GB:M63423
A;Note: the authors translated the codon GGT for residue 6 as Leu, CCC for residue 7 as
R;Hugli, T.E.
J. Biol. Chem. 250, 8293-8301, 1975
A;Title: Human anaphylatoxin (C3a) from the third component of complement.
A;Reference number: A92187; MUID:76069169; PMID:1238393
A;Accession: A92187
A;Molecule type: protein
A;Residues: 672-680,'N',682-699,'Q',701-748 <HUG>
R;Daoudaki, M.E.; Becherer, J.D.; Lambris, J.D.
J. Immunol. 140, 1577-1580, 1988
A;Title: A 34-amino acid peptide of the third component of complement mediates properdin
A;Reference number: A27603; MUID:88154452; PMID:3279119
A;Accession: A27603
A;Molecule type: protein
A;Residues: 1409-1563 <DAO>
R;Hellman, U.; Eggertsen, G.; Engstrom, A.; Sjoquist, J.
Biochem. J. 230, 353-361, 1985
A;Title: Amino acid sequence of the trypsin-generated C3d fragment from human complement
A;Reference number: A23435; MUID:86025442; PMID:3876831
A;Accession: A23435
A;Molecule type: protein
A;Residues: 1002-1012,'E',1014-1303 <HEL>
A;Note: sequence corresponding to residues 1072-1100 was not determined but was taken fr
R;Poznansky, M.C.; Clissold, P.M.; Lachmann, P.J.
J. Immunol. 143, 1254-1258, 1989
A;Title: The difference between human C3F and C3S results from a single amino acid chang
3.
A;Reference number: A45830; MUID:89309808; PMID:2473125

A;Accession: A45830
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1212-1215,'N',1217-1223 <POZ>
A;Note: this is the C3S allele
A;Accession: B45830
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1212-1223 <PO2>
R;Dolmer, K.; Sottrup-Jensen, L.
FEBS Lett. 315, 85-90, 1993
A;Title: Disulfide bridges in human complement component C3b.
A;Reference number: S27041; MUID:93106233; PMID:8416818
A;Contents: annotation; disulfide bonds
C;Comment: The sequence shown is the C3 fast (C3F) allele, which is found mainly in Cauc
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Genetics:
A;Gene: GDB:C3
A;Cross-references: GDB:119044; OMIM:120700
A;Map position: 19p13.3-19p13.3
A;Note: contains 41 exons
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-667/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;23-667,672-1663/Product: complement C3 #status predicted <CC3>
F;23-667,749-1663/Product: C3b #status predicted <C3B>
F;672-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;672-748/Product: C3a anaphylatoxin #status predicted <C3T>
F;749-1663/Product: C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;955-1303/Product: C3dg fragment #status predicted <CDG>
F;955-1001/Product: C3g fragment #status predicted <C3G>
F;1002-1303/Product: C3d fragment #status experimental <C3D>
F;1424-1457/Region: properdin binding
F;85,939/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;559-816,627-662,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;954-955/Cleavage site: Arg-Glu (complement factor I) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status experimental
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 88; DB 1; Length 1663;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
||||||||||||||||
Db 1304 SSKITHRIHWESASLLR 1320

RESULT 2
A27602
complement C3 - rabbit (fragment)
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 15-Dec-1988 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A27602
R;Kusano, M.; Choi, N.H.; Tomita, M.; Yamamoto, K.; Migita, S.; Sekiya, T.; Nishimura, S
Immunol. Invest. 15, 365-378, 1986
A;Title: Nucleotide sequence of cDNA and derived amino acid sequence of rabbit complemen
A;Reference number: A27602; MUID:87006907; PMID:3019881
A;Accession: A27602
A;Molecule type: mRNA
A;Residues: 1-726 <KUS>
A;Cross-references: UNIPROT:P12247; GB:M32434; NID:g164862; PIDN:AAA31190.1; PID:g164863
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;

Query Match 69.3%; Score 61; DB 2; Length 726;
Best Local Similarity 70.6%; Pred. No. 0.017;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
|| : ||| |:||||||
Db 367 SSPVKHRIVWDSASLLR 383

RESULT 3
A82997
hypothetical protein PA5194 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A82997
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A82997
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <STO>
A;Cross-references: UNIPROT:Q9HTZ5; GB:AE004932; GB:AE004091; NID:g9951493; PIDN:AAG0857
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA5194

Query Match 59.1%; Score 52; DB 2; Length 267;
Best Local Similarity 57.1%; Pred. No. 0.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SKITHRIHWESASL 15
:|| | :||: |||
Db 118 AKIAHHLHWQHASL 131

RESULT 4
C3RT
complement C3 precursor - rat
N;Alternate names: 37K phospholipase A2 inhibitory protein
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1992 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: S15764; A54562; A01260; B35979; A35979; PN0567; PN0566; A32281; S08692
R;Misumi, Y.; Sohda, M.; Ikehara, Y.
Nucleic Acids Res. 18, 2178, 1990
A;Title: Nucleotide and deduced amino acid sequence of rat complement C3.
A;Reference number: S15764; MUID:90245672; PMID:2336397
A;Accession: S15764
A;Molecule type: mRNA
A;Residues: 1-1663 <MIS>
A;Cross-references: UNIPROT:P01026; EMBL:X52477; NID:g56953; PIDN:CAA36716.1; PID:g56954
R;Sundstrom, S.A.; Komm, B.S.; Ponce-de-Leon, H.; Yi, Z.; Teuscher, C.; Lyttle, C.R.
J. Biol. Chem. 264, 16941-16947, 1989
A;Title: Estrogen regulation of tissue-specific expression of complement C3.
A;Reference number: A54562; MUID:89380332; PMID:2674144
A;Accession: A54562
A;Status: translation not shown
A;Molecule type: mRNA

```
A;Residues: 'P',1316-1595 <SUN>
A;Cross-references: GB:M29866; NID:g203200; PIDN:AAA40837.1; PID:g554423
R;Jacobs, J.W.; Rubin, J.S.; Hugli, T.E.; Bogardt, R.A.; Mariz, I.K.; Daniels, J.S.; Dau
Biochemistry 17, 5031-5038, 1978
A;Title: Purification, characterization, and amino acid sequence of rat anaphylatoxin (C
A;Reference number: A01260; MUID:79062262; PMID:309768
A;Accession: A01260
A;Molecule type: protein
A;Residues: 671-703,'K',705-720,'KL',723-748 <JAC>
A;Note: three disulfide bonds are present
R;Suwa, Y.; Kudo, I.; Imaizumi, A.; Okada, M.; Kamimura, T.; Suzuki, Y.; Chang, H.W.; Ha
Proc. Natl. Acad. Sci. U.S.A. 87, 2395-2399, 1990
A;Title: Proteinaceous inhibitors of phospholipase A-2 purified from inflammatory sites
A;Reference number: A35979; MUID:90207203; PMID:2320562
A;Accession: B35979
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'X',998-1005 <SUW>
A;Accession: A35979
A;Molecule type: protein
A;Residues: 'X',961-962,'P',964-969 <SU2>
R;Nakagawa, H.; Komorita, N.
Biochem. Biophys. Res. Commun. 194, 1181-1187, 1993
A;Title: Complement component C3-derived neutrophil chemotactic factors purified from ex
A;Reference number: PN0566; MUID:93356786; PMID:8352775
A;Accession: PN0567
A;Molecule type: protein
A;Residues: 568-592 <NAK>
A;Note: amino end of a C3-derived peptide designated exudate neutrophil chemotactic fact
A;Accession: PN0566
A;Molecule type: protein
A;Residues: 671-687 <NA2>
A;Note: amino end of peptide designated neutrophil chemotactic factor 1 and probably ide
R;Kuivanen, P.C.; Capulong, R.B.; Harkins, R.N.; DeSombre, E.R.
Biochem. Biophys. Res. Commun. 158, 898-905, 1989
A;Title: The estrogen-responsive 110K and 74K rat uterine secretory proteins are structu
A;Reference number: A32281; MUID:89149812; PMID:2645873
A;Accession: A32281
A;Molecule type: protein
A;Residues: 25-41 <KUI>
A;Experimental source: 17beta-estradiol-stimulated uterus of immature rat
A;Note: the authors treat this 74K uterine secretory protein, identical as far as sequen
ent
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by prc
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; chemotaxis; complement alternate pathway; complement pathway; g
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666,671-1663/Product: complement C3 #status predicted <CC3>
F;25-666,749-1663/Product: complement C3b #status predicted <C3B>
F;671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;671-748/Product: C3a anaphylatoxin #status experimental <C3T>
F;749-1663/Product: complement C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;1002-1303/Product: C3d fragment #status predicted <C3D>
F;1424-1457/Region: properdin binding
F;558-816,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;939,1617/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

  Query Match             52.3%;  Score 46;  DB 1;  Length 1663;
  Best Local Similarity   58.8%;  Pred. No. 18;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              ||    |: ||| ||||
Db       1304 SSPTVFRLLWESGSLLR 1320


RESULT 5
H83239
pseudouridine synthase RluA PA3246 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83239
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83239
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-211 <STO>
A;Cross-references: UNIPROT:Q9HYZ4; GB:AE004747; GB:AE004091; NID:g9949362; PIDN:AAG0663
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: rluA; PA3246

  Query Match             51.1%;  Score 45;  DB 2;  Length 211;
  Best Local Similarity   40.0%;  Pred. No. 2.6;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLL 16
              ::| ||: ||:: |:
Db         50 ARIVHRLDWETSGLM 64


RESULT 6
F75508
mrr restriction system protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: F75508
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: F75508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <WHI>
A;Cross-references: UNIPROT:Q9RX07; GB:AE001910; GB:AE000513; NID:g6458198; PIDN:AAF1008
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0508
A;Map position: 1

  Query Match             51.1%;  Score 45;  DB 2;  Length 336;
  Best Local Similarity   50.0%;  Pred. No. 4.4;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLLR 17
              ||: ||| |  ::| |
Db         72 SKVRHRIAWACSNLYR 87


RESULT 7
C3MS
complement C3 precursor - mouse
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A92459; B92459; A92460; A93938; A21898; A54561; S16369; S16189; I49563; I49
```

R;Lundwall, A.; Wetsel, R.A.; Domdey, H.; Tack, B.F.; Fey, G.H.
J. Biol. Chem. 259, 13851-13856, 1984
A;Title: Structure of murine complement component C3: I. Nucleotide sequence of cloned c
A;Reference number: A92459; MUID:85054818; PMID:6548745
A;Accession: A92459
A;Molecule type: mRNA
A;Residues: 1-724 <LU1>
A;Cross-references: UNIPROT:P01027
A;Accession: B92459
A;Molecule type: DNA
A;Residues: 1-124 <LU2>
R;Wetsel, R.A.; Lundwall, A.; Davidson, F.; Gibson, T.; Tack, B.F.; Fey, G.H.
J. Biol. Chem. 259, 13857-13862, 1984
A;Title: Structure of murine complement component C3: II. Nucleotide sequence of cloned
A;Reference number: A92460; MUID:85054819; PMID:6094532
A;Accession: A92460
A;Molecule type: mRNA
A;Residues: 671-1663 <WET>
R;Domdey, H.; Wiebauer, K.; Kazmaier, M.; Muller, V.; Odink, K.; Fey, G.
Proc. Natl. Acad. Sci. U.S.A. 79, 7619-7623, 1982
A;Title: Characterization of the mRNA and cloned cDNA specifying the third component of
A;Reference number: A93938; MUID:83117730; PMID:6961437
A;Contents: C3a
A;Accession: A93938
A;Molecule type: mRNA
A;Residues: 671-748 <DOM>
R;Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Lonblad, P.B.; Jones, C.M.; Wierzb
Proc. Natl. Acad. Sci. U.S.A. 82, 9-13, 1985
A;Title: Common evolutionary origin of alpha2-macroglobulin and complement components C3
A;Reference number: A21898; MUID:85113177; PMID:2578664
A;Accession: A21898
A;Molecule type: mRNA
A;Residues: 25-1663 <SOT>
R;Hamada, J.; Cavanaugh, P.G.; Miki, K.; Nicolson, G.L.
Cancer Res. 53, 4418-4423, 1993
A;Title: A paracrine migration-stimulating factor for metastatic tumor cells secreted by
A;Reference number: A54561; MUID:93373334; PMID:8364938
A;Accession: A54561
A;Molecule type: protein
A;Residues: 25-41;749-760 <HAM>
A;Experimental source: migration-stimulating factor purified from medium conditioned by
R;Sato, T.; Hong, M.H.; Jin, C.H.; Ishimi, Y.; Udagawa, N.; Shinki, T.; Abe, E.; Suda, T
FEBS Lett. 285, 21-24, 1991
A;Title: The specific production of the third component of complement by osteoblastic ce
A;Reference number: S16189; MUID:91293304; PMID:2065778
A;Accession: S16369
A;Molecule type: protein
A;Residues: 25-31 <SAT>
A;Accession: S16189
A;Status: preliminary
A;Molecule type: protein
A;Residues: 671-677,'X',679-680 <SA2>
R;Fey, G.; Domdey, H.; Wiebauer, K.; Whitehead, A.S.; Odink, K.
Springer Semin. Immunopathol. 6, 119-147, 1983
A;Title: Structure and expression of the C3 gene.
A;Reference number: I49563; MUID:84045280; PMID:6356427
A;Accession: I49563
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 25-136,'Q',138-240 <FEY>
A;Cross-references: GB:M35659; NID:g192280; PIDN:AAA37339.1; PID:g192281
R;Fey, G.H.; Wiebauer, K.; Domdey, H.
Ann. N. Y. Acad. Sci. 421, 307-312, 1983
A;Title: Amino acid sequences of mouse complement C3 derived from nucleotide sequences c
A;Reference number: I49576; MUID:84201365; PMID:6609661
A;Accession: I49576
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 658-761 <RES>
A;Cross-references: GB:M33032; NID:g192391; PIDN:AAA37378.1; PID:g192392
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.

C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Genetics:
A;Introns: 27/2; 90/3
A;Note: the list of introns may be incomplete
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666,671-1663/Product: complement C3 #status predicted <CC3>
F;25-666,749-1663/Product: C3b #status predicted <C3B>
F;671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;671-748/Product: C3a anaphylatoxin #status predicted <C3T>
F;749-1663/Product: C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;1002-1303/Product: C3d fragment #status predicted <C3D>
F;1424-1457/Region: properdin binding
F;559-816,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;939,1617/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

  Query Match 51.1%; Score 45; DB 1; Length 1663;
  Best Local Similarity 52.9%; Pred. No. 27;
  Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSKITHRIHWESASLLR 17
 || | |: ||: :|||
Db 1304 SSATTFRLLWENGNLLR 1320

RESULT 8
S67037
SMP3 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O3527; protein YOR149c
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S67037; S13750
R;Bordonne, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarassov, I.A.; Winsc
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67032
A;Accession: S67037
A;Molecule type: DNA
A;Residues: 1-516 <BOR>
A;Cross-references: UNIPROT:Q04174; EMBL:Z75057; NID:g1420374; PID:e252038; PID:g1420375
A;Experimental source: strain S288C
R;Irie, K.; Araki, H.; Oshima, Y.
Mol. Gen. Genet. 225, 257-265, 1991
A;Title: Mutations in a Saccharomyces cerevisiae host showing increased holding stabilit
A;Reference number: S13750; MUID:91172125; PMID:2005867
A;Accession: S13750
A;Molecule type: DNA
A;Residues: 1-121,'IK',124-162,'G',164-168,'R',170-278,'L',280-516 <IRI>
A;Cross-references: EMBL:X58121; NID:g4497; PIDN:CAA41123.1; PID:g4498
C;Genetics:
A;Gene: SGD:SMP3
A;Cross-references: SGD:S0005675; MIPS:YOR149c
A;Map position: 15R
C;Keywords: transmembrane protein
F;9-25/Domain: transmembrane #status predicted <TM1>
F;189-205/Domain: transmembrane #status predicted <TM2>
F;215-231/Domain: transmembrane #status predicted <TM3>
F;271-287/Domain: transmembrane #status predicted <TM4>
F;344-360/Domain: transmembrane #status predicted <TM5>

  Query Match 50.0%; Score 44; DB 2; Length 516;
  Best Local Similarity 63.6%; Pred. No. 11;
  Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          6 HRIHWESASLL 16
              :|:||:| |||
Db        207 YRVHWKSFSLL 217
```

RESULT 9
AH0011
ferredoxin-NADP reductase (EC 1.18.1.2) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AH0011
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0011
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-248 <KUR>
A;Cross-references: UNIPROT:Q8ZJK6; GB:AL590842; PIDN:CAC88954.1; PID:g15978201; GSPDB:G
C;Genetics:
A;Gene: fpr
C;Keywords: oxidoreductase

  Query Match 47.7%; Score 42; DB 2; Length 248;
  Best Local Similarity 61.5%; Pred. No. 10;
  Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 SSKITHRIHWESA 13
              | ||||  ||  |
Db          6 SGKITHIEHWTDA 18
```

RESULT 10
C86317
protein T10O22.23 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86317
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86317
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: UNIPROT:Q9LM24; GB:AE005172; NID:g8671774; PIDN:AAF78380.1; GSPDB:GN
C;Genetics:
A;Gene: T10O22.23
A;Map position: 1

  Query Match 47.7%; Score 42; DB 2; Length 280;
  Best Local Similarity 50.0%; Pred. No. 12;
  Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 SSKITHRIHWESASLL 16
              ||  |:|: ||:  ||
Db        122 SSDSTNRLSWENCDLL 137
```

RESULT 11
E82521
hypothetical protein XF2735 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82521
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequer
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-401 <SIM>
A;Cross-references: UNIPROT:Q9P9Y5; GB:AE004080; GB:AE003849; NID:g9107971; PIDN:AAF8552
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohn
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigi
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.,
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; ?
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2735

  Query Match 47.7%; Score 42; DB 2; Length 401;
  Best Local Similarity 45.5%; Pred. No. 18;
  Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          4 ITHRIHWESAS 14
              : ||:||:  |
Db        334 LAHRVHWDEES 344
```

RESULT 12
G75580
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75580
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <WHI>
A;Cross-references: UNIPROT:Q9RYN8; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF1248
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0272
A;Map position: 2
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF0821

  Query Match 47.7%; Score 42; DB 2; Length 474;
  Best Local Similarity 46.2%; Pred. No. 22;
  Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 SSKITHRIHWESA 13
              |:::| |:||  |
Db        396 SARLTSRLHWRPA 408
```

RESULT 13

```
T18946
probable phospholipase activating protein C05C10.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18946; T24252
R;Matthews, P.
submitted to the EMBL Data Library, February 1995
A;Reference number: Z19049
A;Accession: T18946
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-858 <WIL>
A;Cross-references: UNIPROT:Q17647; EMBL:Z48178; PIDN:CAA88206.1; GSPDB:GN00020; CESP:C0
A;Experimental source: clone C05C10
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19863
A;Accession: T24252
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-858 <WI2>
A;Cross-references: EMBL:Z66515; PIDN:CAA91354.1; GSPDB:GN00020; CESP:C05C10.6
A;Experimental source: clone R53
C;Genetics:
A;Gene: CESP:C05C10.6
A;Map position: 2
A;Introns: 15/3; 120/1; 155/3; 407/3; 513/1; 549/1; 593/3; 711/2; 786/3; 821/3

  Query Match             47.7%;  Score 42;  DB 2;  Length 858;
  Best Local Similarity   64.3%;  Pred. No. 42;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          6 HRIHWE--SASLLR 17
              | |||:  |||:||
Db        226 HIIHWDVASASILR 239


RESULT 14
JQ0393
nodulation protein nodA - Azorhizobium caulinodans
N;Alternate names: hypothetical 24.9K protein
C;Species: Azorhizobium caulinodans
A;Note: host Sesbania rostrata
C;Date: 07-Sep-1990 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: JQ0393
R;Goethals, K.; Gao, M.; Tomekpe, K.; Van Montagu, M.; Holsters, M.
Mol. Gen. Genet. 219, 289-298, 1989
A;Title: Common nodABC genes in Nod locus 1 of Azorhizobium caulinodans: nucleotide sequ
A;Reference number: JQ0393; MUID:90136519; PMID:2615763
A;Accession: JQ0393
A;Molecule type: DNA
A;Residues: 1-226 <GOE>
A;Cross-references: UNIPROT:Q07739; GB:L18897; NID:g1293899; PIDN:AAB51162.1; PID:g31029
A;Experimental source: strain ORS571
C;Comment: This is one of the proteins, coded by nodulation genes, that are required for
C;Genetics:
A;Gene: nodA
C;Superfamily: nodulation protein nodA
C;Keywords: nodulation

  Query Match             46.6%;  Score 41;  DB 1;  Length 226;
  Best Local Similarity   63.6%;  Pred. No. 14;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWES 12
              ||:| |: |||
Db         33 SKVTWRVAWES 43


RESULT 15
AI3289
hypothetical cytosolic protein BMEI0303 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AI3289
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AI3289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229 <KUR>
A;Cross-references: UNIPROT:Q8YIY6; UNIPROT:Q8FYX0; GB:AE008917; PIDN:AAL51484.1; PID:g1
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0303
A;Map position: I
C;Superfamily: Rickettsia prowazekii hypothetical protein RP073

  Query Match             46.6%;  Score 41;  DB 2;  Length 229;
  Best Local Similarity   53.8%;  Pred. No. 14;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 KITHRIHWESASL 15
              :| :| || ||:|
Db        136 QIRNRTHWNSANL 148


Search completed: June  1, 2005, 09:34:51
Job time : 40 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 1, 2005, 09:08:01 ; Search time 167 Seconds
(without alignments)
52.128 Million cell updates/sec

Title: US-09-845-736-1
Perfect score: 88
Sequence: 1 SSKITHRIHWESASLLR 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	1663	1	CO3_HUMAN	P01024 homo sapien
2	61	69.3	154	2	Q29289	Q29289 sus scrofa
3	61	69.3	726	1	CO3_RABIT	P12247 oryctolagus
4	61	69.3	1661	2	Q9GKP1	Q9gkp1 sus scrofa
5	60	68.2	167	2	Q9N0M4	Q9n0m4 cervus nipp
6	60	68.2	349	2	O46544	O46544 ovis aries
7	52	59.1	267	2	Q9HTZ5	Q9htz5 pseudomonas
8	49	55.7	229	2	Q6FYQ3	Q6fyq3 bartonella
9	46	52.3	441	2	Q8T3J9	Q8t3j9 drosophila
10	46	52.3	1663	1	CO3_RAT	P01026 rattus norv
11	45	51.1	75	2	Q87EL5	Q87el5 xylella fas
12	45	51.1	211	2	Q9HYZ4	Q9hyz4 pseudomonas
13	45	51.1	336	2	Q9RX07	Q9rx07 deinococcus
14	45	51.1	422	2	Q6GLI2	Q6gli2 xenopus tro
15	45	51.1	440	2	Q6DJB7	Q6djb7 xenopus tro
16	45	51.1	1663	1	CO3_MOUSE	P01027 mus musculu
17	45	51.1	1663	2	Q80XP1	Q80xp1 mus musculu
18	44	50.0	196	2	Q7PK14	Q7pk14 anopheles g
19	44	50.0	422	2	Q7PK13	Q7pk13 anopheles g
20	44	50.0	516	1	SMP3_YEAST	Q04174 saccharomyc
21	44	50.0	545	2	Q84MM3	Q84mm3 vigna ungui
22	44	50.0	1470	2	Q8I266	Q8i266 plasmodium
23	43	48.9	75	2	Q9GMH7	Q9gmh7 macaca fasc
24	43	48.9	384	2	Q7VRT2	Q7vrt2 candidatus
25	42	47.7	173	2	Q6MMC3	Q6mmc3 bdellovibri
26	42	47.7	232	2	Q6G5P3	Q6g5p3 bartonella
27	42	47.7	248	2	Q66G98	Q66g98 yersinia ps
28	42	47.7	248	2	Q74Y75	Q74y75 yersinia pe
29	42	47.7	248	2	Q8ZJK6	Q8zjk6 yersinia pe
30	42	47.7	280	2	Q9LM24	Q9lm24 arabidopsis
31	42	47.7	338	2	Q8PF47	Q8pf47 xanthomonas
32	42	47.7	358	2	Q8IYM1	Q8iym1 homo sapien
33	42	47.7	358	2	Q96LL0	Q96ll0 homo sapien
34	42	47.7	360	2	Q9LPP7	Q9lpp7 arabidopsis
35	42	47.7	401	2	Q879V8	Q879v8 xylella fas
36	42	47.7	401	2	Q9P9Y5	Q9p9y5 xylella fas
37	42	47.7	407	2	Q8SYY7	Q8syy7 drosophila
38	42	47.7	414	2	Q7UNU7	Q7unu7 rhodopirell
39	42	47.7	474	2	Q9RYN8	Q9ryn8 deinococcus
40	42	47.7	493	2	Q6ESY1	Q6esy1 oryza sativ
41	42	47.7	858	2	Q17647	Q17647 caenorhabdi
42	42	47.7	860	2	Q95NM4	Q95nm4 caenorhabdi
43	42	47.7	1254	2	Q7S4X1	Q7s4x1 neurospora
44	41.5	47.2	381	2	Q8RTQ7	Q8rtq7 thermodesul
45	41.5	47.2	382	2	Q93EV7	Q93ev7 thermodesul

ALIGNMENTS

RESULT 1
CO3_HUMAN
ID CO3_HUMAN STANDARD; PRT; 1663 AA.
AC P01024;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN Name=C3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85140166; PubMed=2579379;
RA de Bruijn M.H.L., Fey G.H.;
RT "Human complement component C3: cDNA coding sequence and derived
RT primary structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:708-712(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS GLY-102; PRO-314; LYS-863; ASP-1224
RP AND THR-1367.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 672-748.
RX MEDLINE=76069169; PubMed=1238393;
RA Hugli T.E.;
RT "Human anaphylatoxin (C3a) from the third component of complement.
RT Primary structure.";
RL J. Biol. Chem. 250:8293-8301(1975).
RN [4]
RP SEQUENCE OF 955-966, AND SUBUNITS.
RC TISSUE=Serum;
RX MEDLINE=95293954; PubMed=7539791; DOI=10.1074/jbc.270.23.13645;
RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
RA Stigbrand T., Gleich G.J., Sottrup-Jensen L.;
RT "Identification of angiotensinogen and complement C3dg as novel
RT proteins binding the proform of eosinophil major basic protein in
RT human pregnancy serum and plasma.";
RL J. Biol. Chem. 270:13645-13651(1995).
RN [5]
RP SEQUENCE OF 988-1036.
RX MEDLINE=82174534; PubMed=6175959;
RA Thomas M.L., Janatova J., Gray W.R., Tack B.F.;
RT "Third component of human complement: localization of the internal
RT thiolester bond.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1054-1058(1982).
RN [6]
RP SEQUENCE OF 1409-1563.

RX MEDLINE=88154452; PubMed=3279119;
RA Daoudaki M.E., Becherer J.D., Lambris J.D.;
RT "A 34-amino acid peptide of the third component of complement mediates
RT properdin binding.";
RL J. Immunol. 140:1577-1580(1988).
RN [7]
RP STRUCTURE BY NMR OF C3A.
RX MEDLINE=88276894; PubMed=3260670;
RA Nettesheim D.G., Edalji R.P., Mollison K.W., Greer J.,
RA Zuiderweg E.R.P.;
RT "Secondary structure of complement component C3a anaphylatoxin in
RT solution as determined by NMR spectroscopy: differences between
RT crystal and solution conformations.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5036-5040(1988).
RN [8]
RP MUTAGENESIS OF THIOESTER BOND REGION.
RX MEDLINE=92250565; PubMed=1577777;
RA Isaac L., Isenman D.E.;
RT "Structural requirements for thioester bond formation in human
RT complement component C3. Reassessment of the role of thioester bond
RT integrity on the conformation of C3.";
RL J. Biol. Chem. 267:10062-10069(1992).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=93106233; PubMed=8416818; DOI=10.1016/0014-5793(93)81139-Q;
RA Dolmer K., Sottrup-Jensen L.;
RT "Disulfide bridges in human complement component C3b.";
RL FEBS Lett. 315:85-90(1993).
RN [10]
RP CARBOHYDRATE-LINKAGE SITE ASN-85.
RX MEDLINE=22660472; PubMed=12754519; DOI=10.1038/nbt827;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 996-1303.
RX MEDLINE=98259089; PubMed=9596584; DOI=10.1126/science.280.5367.1277;
RA Nagar B., Jones R.G., Diefenbach R.J., Isenman D.E., Rini J.M.;
RT "X-ray crystal structure of C3d: a C3 fragment and ligand for
RT complement receptor 2.";
RL Science 280:1277-1281(1998).
RN [12]
RP VARIANT C3F/S.
RX MEDLINE=89309808; PubMed=2473125;
RA Poznansky M.C., Clissold P.M., Lachmann P.J.;
RT "The difference between human C3F and C3S results from a single amino
RT acid change from an asparagine to an aspartate residue at position
RT 1216 on the alpha-chain of the complement component, C3.";
RL J. Immunol. 143:1254-1258(1989).
RN [13]
RP ERRATUM (RETRACTION).
RX MEDLINE=90063087; PubMed=2584723;
RA Poznansky M.C., Clissold P.M., Lachmann P.J.;
RL J. Immunol. 143:3860-3862(1989).
RN [14]
RP VARIANTS GLY-102 AND PRO-314.
RX MEDLINE=91011240; PubMed=1976733;
RA Botto M., Yong Fong K., So A.K., Koch C., Walport M.J.;
RT "Molecular basis of polymorphisms of human complement component C3.";
RL J. Exp. Med. 172:1011-1017(1990).
RN [15]
RP VARIANT ASN-549.
RX MEDLINE=95050640; PubMed=7961791;
RA Singer L., Whitehead W.T., Akama H., Katz Y., Fishelson Z.,
RA Wetsel R.A.;
RT "Inherited human complement C3 deficiency. An amino acid substitution
RT in the beta-chain (Asp549 to Asn) impairs C3 secretion.";
RL J. Biol. Chem. 269:28494-28499(1994).
RN [16]
RP VARIANT GLN-1320.
RA Watanabe Y., Matsui N., Yan K., Nishimukai H., Tokunaga K., Juji T.,
RA Kobayashi N., Kohsaka T.;
RT "A novel C3 allotype C3'F02'has an amino acid substitution that may
RT inhibit iC3b synthesis and cause C3-hypocomplementemia.";
RL Mol. Immunol. 30:62-62(1993).
CC -!- FUNCTION: C3 plays a central role in the activation of the
CC complement system. Its processing by C3 convertase is the central
CC reaction in both classical and alternative complement pathways.
CC After activation C3b can bind covalently, via its reactive
CC thiolester, to cell surface carbohydrates or immune aggregates.
CC -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC C3a anaphylatoxin is a mediator of local inflammatory process. It
CC induces the contraction of smooth muscle, increases vascular
CC permeability and causes histamine release from mast cells and
CC basophilic leukocytes.
CC -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC residues, forming two chains, beta and alpha, linked by a
CC disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC + alpha' chain). During pregnancy, C3dg exists as a complex
CC (probably a 2:2:2 heterohexamer) with AGT and the proform of PRG2.
CC -!- PTM: C3b is rapidly split in two positions by factor I and a
CC cofactor to form iC3b (inactivated C3b) and C3f which is released.
CC Then iC3b is slowly cleaved (possibly by factor I) to form C3c and
CC C3dg. Other proteases produce other fragments such as C3d or C3g.
CC -!- POLYMORPHISM: There are two alleles: C3S (C3 slow), the most
CC common allele in all races and C3F (C3 fast), relatively frequent
CC in Caucasoids, less common in Black Americans, extremely rare in
CC Orientals.
CC -!- DISEASE: Defects in C3 are the cause of C3 deficiency
CC [MIM:120700]. It can result in susceptibility to pyogenic
CC infection.
CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC -!- SIMILARITY: Contains 1 NTR domain.
CC ---

CC ---
DR EMBL; K02765; AAA85332.1; -.
DR EMBL; AY513239; AAR89906.1; -.
DR PIR; A94065; C3HU.
DR PDB; 1C3D; X-ray; @=-.
DR PDB; 1GHQ; X-ray; A=994-1300.
DR GlycoSuiteDB; P01024; -.
DR SWISS-2DPAGE; P01024; HUMAN.
DR Siena-2DPAGE; P01024; -.
DR Genew; HGNC:1318; C3.
DR MIM; 120700; -.
DR GO; GO:0005102; F:receptor binding; TAS.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002890; A2M_N.
DR InterPro; IPR009048; AM_receptor_bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxn.
DR InterPro; IPR008964; Invasin_intimin.
DR InterPro; IPR001599; MacrogloblnA2.
DR InterPro; IPR001134; Netrin_C.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.

```
DR   PROSITE; PS50189; NTR; 1.
KW   3D-structure; Complement alternate pathway; Complement pathway;

  Query Match             100.0%;  Score 88;  DB 1;  Length 1663;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-06;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |||||||||||||||||
Db       1304 SSKITHRIHWESASLLR 1320


RESULT 2
Q29289
ID   Q29289      PRELIMINARY;      PRT;   154 AA.
AC   Q29289;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Complement C3 (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Small intestine;
RX   MEDLINE=96327607; PubMed=8672129;
RA   Winteroe A.K., Fredholm M., Davies W.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Mamm. Genome 7:509-517(1996).
DR   EMBL; F14640; CAA23173.1; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR008930; Terp_cyc_toroid.
FT   NON_TER       1      1
FT   NON_TER     154    154
SQ   SEQUENCE   154 AA;  17440 MW;  6DC7661C1253ED45 CRC64;

  Query Match             69.3%;  Score 61;  DB 2;  Length 154;
  Best Local Similarity   70.6%;  Pred. No. 0.026;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |: : ||| ||||||||
Db         97 SAPVRHRILWESASLLR 113


RESULT 3
CO3_RABIT
ID   CO3_RABIT      STANDARD;      PRT;   726 AA.
AC   P12247;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Complement C3 alpha chain (Fragment).
GN   Name=C3;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87006907; PubMed=3019881;
RA   Kusano M., Choi N.H., Tomita M., Yamamoto K., Migita S., Sekiya T.,
RA   Nishimura S.;
RT   "Nucleotide sequence of cDNA and derived amino acid sequence of rabbit
RT   complement component C3 alpha-chain.";
RL   Immunol. Invest. 15:365-378(1986).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32434; AAA31190.1; -.
DR   PIR; A27602; A27602.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; PARTIAL.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; PARTIAL.
DR   PROSITE; PS50189; NTR; 1.
KW   Complement alternate pathway; Complement pathway; Glycoprotein;
KW   Inflammatory response; Plasma; Thioester bond.
FT   NON_TER       1      1
FT   CHAIN        <1    726       Complement C3 alpha chain.
FT   DOMAIN      581    724       NTR.
FT   CROSSLNK     73     76       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD      2      2       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    680    680       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   726 AA;  81844 MW;  F4B4C35D461300E9 CRC64;

  Query Match             69.3%;  Score 61;  DB 1;  Length 726;
  Best Local Similarity   70.6%;  Pred. No. 0.14;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              || : ||| |:||||||
Db        367 SSPVKHRIVWDSASLLR 383


RESULT 4
Q9GKP1
ID   Q9GKP1      PRELIMINARY;      PRT;  1661 AA.
AC   Q9GKP1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Complement component C3 (Complement C3).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=21313047; PubMed=11419349;
RA   Wimmers K., Mekchay S., Ponsuksili S., Hardge T., Yerle M.,
RA   Schellander K.;
RT   "Polymorphic sites in exon 15 and 30 of the porcine C3 gene.";
RL   Anim. Genet. 32:46-47(2001).
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Wimmers K., Ponsuksili S., Schmoll F., Schellander K.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22444329; PubMed=12557058;
RA   Wimmers K., Mekchay S., Schellander K., Ponsuksili S.;
RT   "Molecular characterization of the pig C3 gene and its association
RT   with complement activity.";
RL   Immunogenetics 54:714-724(2003).
DR   EMBL; AF154933; AAG40565.1; -.
DR   EMBL; AJ494748; CAD38823.2; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0006956; P:complement activation; IEA.
DR   GO; GO:0006954; P:inflammatory response; IEA.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
SQ   SEQUENCE   1661 AA;  186805 MW;  4899D0914BE3310C CRC64;

  Query Match             69.3%;  Score 61;  DB 2;  Length 1661;
  Best Local Similarity   70.6%;  Pred. No. 0.35;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              |: : ||| ||||||||
Db       1302 SAPVRHRILWESASLLR 1318


RESULT 5
Q9N0M4
ID   Q9N0M4      PRELIMINARY;      PRT;   167 AA.
AC   Q9N0M4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Complement C3 alpha chain (Fragment).
OS   Cervus nippon (Sika deer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC   Cervinae; Cervus.
OX   NCBI_TaxID=9863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Jiang Y., Sun L.G., Yu Y.L.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF264631; AAF73464.1; -.
DR   HSSP; P01024; 1C3D.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR008930; Terp_cyc_toroid.
FT   NON_TER       1      1
SQ   SEQUENCE   167 AA;  18671 MW;  12BFE0798290DFA7 CRC64;

  Query Match             68.2%;  Score 60;  DB 2;  Length 167;
  Best Local Similarity   70.6%;  Pred. No. 0.042;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              :| : ||| ||||||||
Db         47 NSLVKHRILWESASLLR 63


RESULT 6
O46544
ID   O46544      PRELIMINARY;      PRT;   349 AA.
AC   O46544;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Complement component C3 (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=white alpine; TISSUE=Liver;
RX   MEDLINE=98309471; PubMed=9647256;
RA   Hein W.R., Dudler L., Marston W.L., Landsverk T., Young A.J.,
RA   Avila D.;
RT   "Ubiquitination and dimerization of complement receptor type 2 on
RT   sheep B cells.";
RL   J. Immunol. 161:458-466(1998).
DR   EMBL; AF038130; AAB92374.2; -.
DR   HSSP; P01026; 1QQF.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
FT   NON_TER       1      1
FT   NON_TER     349    349
SQ   SEQUENCE   349 AA;  39679 MW;  70C2023E42ED5EE3 CRC64;

  Query Match             68.2%;  Score 60;  DB 2;  Length 349;
  Best Local Similarity   70.6%;  Pred. No. 0.094;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              :| : ||| ||||||||
Db        328 NSLVKHRILWESASLLR 344


RESULT 7
Q9HTZ5
ID   Q9HTZ5      PRELIMINARY;      PRT;   267 AA.
AC   Q9HTZ5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PA5194;
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
```

```
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004932; AAG08579.1; -.
DR   PIR; A82997; A82997.
DR   InterPro; IPR008934; AcPase_VanPerase.
DR   InterPro; IPR000326; Pesterase_PA_PTP.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   267 AA;  30527 MW;  57CD9D2319B6AD7D CRC64;

  Query Match             59.1%;  Score 52;  DB 2;  Length 267;
  Best Local Similarity   57.1%;  Pred. No. 1.6;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASL 15
              :|| | :||: |||
Db        118 AKIAHHLHWQHASL 131


RESULT 8
Q6FYQ3
ID   Q6FYQ3      PRELIMINARY;      PRT;   229 AA.
AC   Q6FYQ3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=BQ11780;
OS   Bartonella quintana (Rochalimaea quintana).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bartonellaceae; Bartonella.
OX   NCBI_TaxID=803;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Toulouse;
RX   PubMed=15210978; DOI=10.1073/pnas.0305659101;
RA   Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA   Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
RA   La Scola B., Holmberg M., Andersson S.G.E.;
RT   "The louse-borne human pathogen Bartonella quintana is a genomic
RT   derivative of the zoonotic agent Bartonella henselae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
DR   EMBL; BX897700; CAF26637.1; -.
DR   InterPro; IPR010421; DUF1013.
DR   Pfam; PF06242; DUF1013; 1.
KW   Complete proteome.
SQ   SEQUENCE   229 AA;  25525 MW;  71C34119CE1A6A6F CRC64;

  Query Match             55.7%;  Score 49;  DB 2;  Length 229;
  Best Local Similarity   57.1%;  Pred. No. 4.4;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 KITHRIHWESASLL 16
              :| || || ||:|:
Db        136 QIRHRTHWNSANLV 149


RESULT 9
Q8T3J9
ID   Q8T3J9      PRELIMINARY;      PRT;   441 AA.
AC   Q8T3J9; Q9VLX7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   AT11889p (CG7196-PA).
GN   ORFNames=CG7196;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [5]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY094997; AAM11325.1; -.
DR   EMBL; AE003618; AAF52552.2; -.
DR   FlyBase; FBgn0031944; CG7196.
SQ   SEQUENCE   441 AA;  52125 MW;  847067D8FA3A3A16 CRC64;

  Query Match             52.3%;  Score 46;  DB 2;  Length 441;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 KITHRIHWESASLL 16
              |: |: ||   |||
Db         20 KVVHKNHWRQVSLL 33


RESULT 10
CO3_RAT
ID   CO3_RAT        STANDARD;      PRT;  1663 AA.
AC   P01026;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Complement C3 precursor [Contains: C3a anaphylatoxin].
GN   Name=C3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=90245672; PubMed=2336397;
RA   Misumi Y., Sohda M., Ikehara Y.;
RT   "Nucleotide and deduced amino acid sequence of rat complement C3.";
RL   Nucleic Acids Res. 18:2178-2178(1990).
RN   [2]
RP   SEQUENCE OF 671-748.
RX   MEDLINE=79062262; PubMed=309768;
RA   Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA   Daniels J.S., Daughaday W.H., Bradshaw R.A.;
RT   "Purification, characterization, and amino acid sequence of rat
RT   anaphylatoxin (C3a).";
RL   Biochemistry 17:5031-5038(1978).
RN   [3]
RP   SEQUENCE OF 1316-1595 FROM N.A.
RX   MEDLINE=89380332; PubMed=2674144;
RA   Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA   Lyttle C.R.;
RT   "Estrogen regulation of tissue-specific expression of complement C3.";
RL   J. Biol. Chem. 264:16941-16947(1989).
CC   -!- FUNCTION: C3 plays a central role in the activation of the
CC       complement system. Its processing by C3 convertase is the central
CC       reaction in both classical and alternative complement pathways.
CC       After activation C3b can bind covalently, via its reactive
CC       thiolester, to cell surface carbohydrates or immune aggregates.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC       C3a anaphylatoxin is a mediator of local inflammatory process. It
CC       induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC       + alpha' chain).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52477; CAA36716.1; -.
DR   EMBL; M29866; AAA40837.1; ALT_SEQ.
DR   PIR; S15764; C3RT.
DR   PDB; 1QQF; X-ray; A=1010-1286.
DR   PDB; 1QSJ; X-ray; A/B/C/D=1010-1286.
DR   RGD; 2232; C3.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR009048; AM_receptor_bind.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR008964; Invasin_intimin.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   InterPro; IPR008930; Terp_cyc_toroid.
DR   InterPro; IPR008993; TIMP_like.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR   PROSITE; PS50189; NTR; 1.
KW   3D-structure; Complement alternate pathway; Complement pathway;
KW   Direct protein sequencing; Glycoprotein; Inflammatory response;
KW   Plasma; Signal; Thioester bond.
FT   SIGNAL       1     24
FT   CHAIN       25   1663       Complement C3.
FT   CHAIN       25    666       Complement C3 beta chain.
FT   CHAIN      671   1663       Complement C3 alpha chain.
FT   PEPTIDE    671    748       C3a anaphylatoxin.
FT   CHAIN      749   1663       Complement C3b alpha' chain.
FT   DOMAIN     693    728       Anaphylatoxin-like.
FT   DOMAIN    1518   1661       NTR.
FT   SITE       748    749       Cleavage (by C3 convertase).
FT   DISULFID   558    816       Interchain (By similarity).
FT   DISULFID   626    661       By similarity.
FT   DISULFID   693    720       By similarity.
FT   DISULFID   694    727       By similarity.
FT   DISULFID   707    728       By similarity.
FT   DISULFID   873   1513       By similarity.
FT   DISULFID  1101   1158       By similarity.
FT   DISULFID  1358   1489       By similarity.
FT   DISULFID  1389   1458       By similarity.
FT   DISULFID  1506   1511       By similarity.
FT   DISULFID  1518   1590       By similarity.
FT   DISULFID  1537   1661       By similarity.
FT   CROSSLNK  1010   1013       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD   939    939       N-linked (GlcNAc. . .) (Probable).
```

```
FT   CARBOHYD    1617   1617       N-linked (GlcNAc. . .) (Probable).
FT   CONFLICT     721    722       LK -> KL (in Ref. 2).
FT   TURN        1011   1012
FT   HELIX       1013   1031
FT   TURN        1032   1032
FT   HELIX       1034   1037
FT   HELIX       1039   1041
FT   HELIX       1042   1057
FT   TURN        1058   1059
FT   STRAND      1060   1060
FT   TURN        1062   1063
FT   STRAND      1066   1066
FT   TURN        1070   1071
FT   HELIX       1076   1089
FT   TURN        1090   1092
FT   HELIX       1097   1111
FT   STRAND      1112   1112
FT   TURN        1114   1115
FT   STRAND      1118   1118
FT   HELIX       1127   1134
FT   TURN        1137   1138
FT   HELIX       1139   1158
FT   TURN        1159   1161
FT   TURN        1163   1164
FT   HELIX       1165   1180
FT   TURN        1181   1182
FT   HELIX       1186   1198
FT   TURN        1199   1200
FT   TURN        1204   1205
FT   HELIX       1206   1213
FT   STRAND      1215   1215
FT   TURN        1216   1218
FT   STRAND      1219   1219
FT   TURN        1223   1224
FT   HELIX       1226   1242
FT   TURN        1243   1244
FT   TURN        1246   1247
FT   HELIX       1249   1258
FT   TURN        1263   1264
FT   TURN        1266   1267
FT   HELIX       1269   1285
SQ   SEQUENCE   1663 AA;  186460 MW;  2F87CCB143CDD4BC CRC64;

  Query Match             52.3%;  Score 46;  DB 1;  Length 1663;
  Best Local Similarity   58.8%;  Pred. No. 1.2e+02;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 SSKITHRIHWESASLLR 17
              ||    |: ||| ||||
Db       1304 SSPTVFRLLWESGSLLR 1320


RESULT 11
Q87EL5
ID   Q87EL5      PRELIMINARY;      PRT;    75 AA.
AC   Q87EL5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PD0289;
OS   Xylella fastidiosa (strain Temecula1 / ATCC 700964).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=183190;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22421331; PubMed=12533478;
RX   DOI=10.1128/JB.185.3.1018-1026.2003;
RA   Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA   Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA   Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA   Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA   Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA   Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA   Marino C.L., Giglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA   Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA   da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA   Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA   de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA   Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA   Kitajima J.P.;
RT   "Comparative analyses of the complete genome sequences of Pierce's
RT   disease and citrus variegated chlorosis strains of Xylella
RT   fastidiosa.";
RL   J. Bacteriol. 185:1018-1026(2003).
DR   EMBL; AE012554; AAO28174.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   75 AA;  8640 MW;  83E54A9F163B32B7 CRC64;

  Query Match             51.1%;  Score 45;  DB 2;  Length 75;
  Best Local Similarity   43.8%;  Pred. No. 6.1;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLLR 17
              | :| |:||   |:|:
Db         54 SSLTFRVHWLQVSILQ 69


RESULT 12
Q9HYZ4
ID   Q9HYZ4      PRELIMINARY;      PRT;   211 AA.
AC   Q9HYZ4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Pseudouridine synthase RluA.
GN   Name=rluA; OrderedLocusNames=PA3246;
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC       5'-phosphate + H(2)O.
CC   -!- SIMILARITY: Belongs to the pseudouridine synthase rluA family.
DR   EMBL; AE004747; AAG06634.1; -.
DR   PIR; H83239; H83239.
DR   HSSP; Q8X9F0; 1PRZ.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0009982; F:pseudouridine synthase activity; IEA.
DR   GO; GO:0004730; F:pseudouridylate synthase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0006396; P:RNA processing; IEA.
DR   InterPro; IPR006145; PseudoU_synth.
DR   InterPro; IPR006224; Rlu_synth.
DR   Pfam; PF00849; PseudoU_synth_2; 1.
DR   ProDom; PD001819; PseudoU_synth; 1.
DR   PROSITE; PS01129; PSI_RLU; 1.
KW   Complete proteome; Lyase.
SQ   SEQUENCE   211 AA;  24339 MW;  D333B20FCEA55A94 CRC64;
```

```
  Query Match             51.1%;  Score 45;  DB 2;  Length 211;
  Best Local Similarity   40.0%;  Pred. No. 19;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLL 16
              ::| ||: ||:: |:
Db         50 ARIVHRLDWETSGLM 64


RESULT 13
Q9RX07
ID   Q9RX07      PRELIMINARY;      PRT;   336 AA.
AC   Q9RX07;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mrr restriction system protein.
GN   OrderedLocusNames=DR0508;
OS   Deinococcus radiodurans.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX   MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001910; AAF10088.1; -.
DR   PIR; F75508; F75508.
DR   TIGR; DR0508; -.
DR   InterPro; IPR007560; Mrr_cat.
DR   Pfam; PF04471; Mrr_cat; 1.
KW   Complete proteome.
SQ   SEQUENCE   336 AA;  37335 MW;  E978C50EC4BBC17B CRC64;

  Query Match             51.1%;  Score 45;  DB 2;  Length 336;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SKITHRIHWESASLLR 17
              ||: ||| |  ::| |
Db         72 SKVRHRIAWACSNLYR 87


RESULT 14
Q6GLI2
ID   Q6GLI2      PRELIMINARY;      PRT;   422 AA.
AC   Q6GLI2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC69416 protein.
GN   Name=MGC69416;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; BC074508; AAH74508.1; -.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   422 AA;  48355 MW;  FF99B876238FA1D1 CRC64;

  Query Match             51.1%;  Score 45;  DB 2;  Length 422;
  Best Local Similarity   50.0%;  Pred. No. 41;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 ITHRIHWESASLLR 17
              ::|| |:|: :|||
Db        117 LSHRFHYENPTLLR 130


RESULT 15
Q6DJB7
ID   Q6DJB7      PRELIMINARY;      PRT;   440 AA.
AC   Q6DJB7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC88881 protein.
GN   Name=MGC88881;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8364;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
```

```
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Whole body;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; BC075265; AAH75265.1; -.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   440 AA;  50228 MW;  0F0AF12772CFA9D9 CRC64;

  Query Match             51.1%;  Score 45;  DB 2;  Length 440;
  Best Local Similarity   50.0%;  Pred. No. 43;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 ITHRIHWESASLLR 17
              ::|| |:|: :|||
Db        137 LSHRFHYENPTLLR 150


Search completed: June  1, 2005, 09:34:08
Job time : 170 secs
```